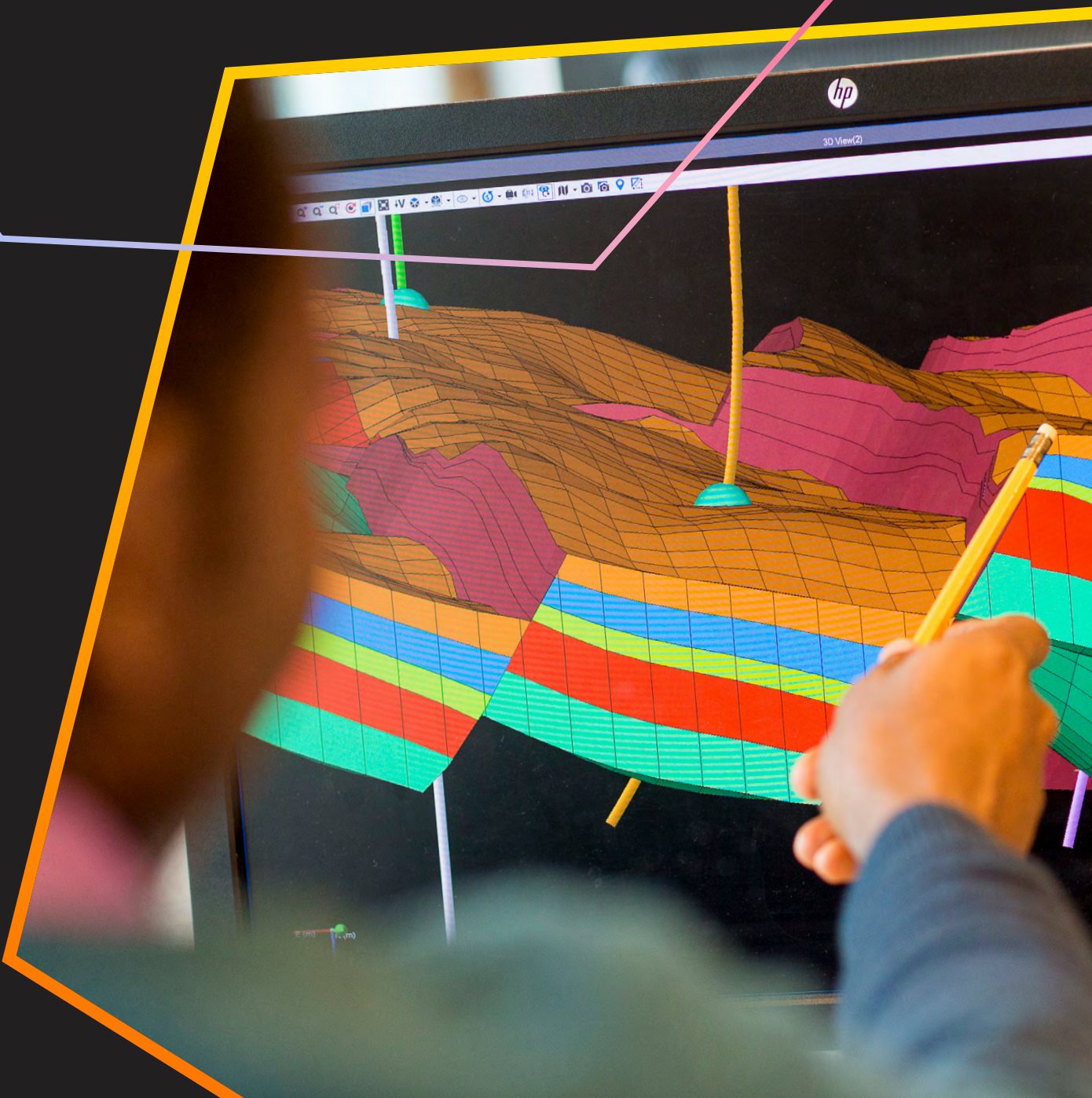


Data-centric engineering and AI skills for engineers

May 2026



Data-centric engineering and AI skills for engineers

A joint report by the Royal Academy of
Engineering, The Alan Turing Institute
and Lloyd's Register Foundation

ISBN:

978-1-909327-58-0

Available to download from:

<https://www.raeng.org.uk/data-centric-engineering-ai-skills-report>

Authors

Dr Rhys Morgan (Royal Academy of Engineering),
Dr Gabin Kayumbi (The Alan Turing Institute)

Acknowledgements

Steering Group:

Chair: Professor Adam Sobey (The Alan Turing Institute,
University of Southampton)

Members:

Professor Tim Broyd FREng CEng (University College London)
Professor Elizabeth Cross (University of Sheffield)
Lydia Amarquaye CEng (The Institution of Mechanical Engineers)
Professor John Chudley CEng (Engineering Council)
Poppy Harrison CEng (AtkinsRéalis)
Katy Henderson (The Alan Turing Institute)
Professor Claire Lucas CEng (Aston University)
Professor Omar Matar FREng CEng (Imperial College London)
Dr Vera Matser (The Alan Turing Institute)
Dr Thomas Popham (University of Warwick)
Dr Annalisa Riccardi (University of Strathclyde)
David Short FREng CEng (BAE Systems)
Dr Tim Slingsby (Lloyd's Register Foundation)

The authors would like to thank the many academics who
contributed to the survey and qualitative interviews.

Contents

Executive summary	1
Recommendations	2
1. Introduction	5
1.1 Background	5
1.2 Purpose and scope	6
1.3 Project stages	7
2. Methodology	8
3. Data-centric engineering learning outcomes	10
3.1 Tiered learning structure	10
3.2 'Modified' Bloom's taxonomy	11
3.3 Proposed DCE learning outcomes	11
1. Foundational statistics	13
2. Advanced statistics	14
3. Data engineering and management	15
4. Data analytics and visualisation	16
5. Data governance, ethics and communication	17
6. ML and AI for data analytics	18
7. Additional AI skills	19
3.4 Obsolescence of the learning outcomes	20
4. HE capacity to deliver the learning outcomes	22
4.1 Quantitative analysis	22
4.2 Qualitative: Thematic analysis	25
5. Discussion and recommendations	29
5.1 Discussion	29
5.2 Recommendations	31
Annex A: Methodology	34
Annex B: Results	36

Executive summary

Data has been a core component of engineering since its early beginnings, supporting physical laws developed over the past several hundred years that are used in engineering today. However, engineering is now evolving to a point where knowledge and deployment of empirical-based laws are no longer sufficient for engineers working in a world of increasingly interconnected and complex systems and rapidly emerging new technology.

Data-centric engineering (DCE) is emerging as a new discipline to address this, fusing the fundamentals of science (and social science) with data to explain complex engineering products, systems and processes. Alongside this, machine learning (ML) and artificial intelligence (AI) have become powerful tools to manipulate, analyse and interpret large datasets. In this new paradigm, engineers need a new suite of knowledge, skills and behaviours to work effectively.

As industries shift toward data-driven operations, universities play a critical role in preparing graduates with the skills required to enable this transformation. This report sets out the learning outcomes for engineering undergraduates in UK higher education institutions aligned to DCE and the use of AI tools for engineering, drawing on the power of data to optimise all aspects of the engineering lifecycle.

We identify fifty learning outcomes across seven thematic areas:

- Foundational statistics
- Advanced Statistics
- Data engineering and management
- Data analytics
- Data governance
- ML and AI for data analytics
- Additional AI skills for engineering

Having identified the thematic areas, we analysed the extent to which the learning outcomes are currently embedded in UK higher education engineering programmes across all disciplines, exploring the current curriculum landscape and pedagogical practices. We found that certain aspects of data-handling literacy (foundational statistics and analytics) are well embedded across UK engineering higher education. However, more advanced and infrastructural aspects of DCE, such as data engineering and advanced AI, are often concentrated in certain disciplines or, are often only explored in depth by subsets of students in certain universities who choose particular data-focused

elective modules in latter stages of their degree or who pursue data-intensive, final year, individual or team-based projects.

The transition to data-driven engineering practice demands both curricular reform and a systemic shift in pedagogical culture. However, there is broad consensus across the academic community that the curriculum is already too full, and therefore DCE must be embedded as an integrative capability across the engineering curriculum, rather than positioned as an additional, isolated subject. This means greater use of data engineering and AI through lab exercises, and in problem-based and project-based work. As this integration is considered, the academic community must come to terms with the fact that the advanced tools associated with data engineering, particularly AI, are evolving at such a rapid pace that engineering students will be using them without fully understanding how they work. However, this also underlines the importance of students understanding the fundamentals of engineering disciplines, so that they can intuit whether the information presented by AI and other data tools is correct.

Institutional capacity is identified as a key driver of DCE maturity. Challenges for greater embedding of DCE and AI include uneven staff capability and limited resources for adoption across higher education. Staff development, pedagogical support, project-based application and cross-sector collaboration will all be needed to embed further use of DCE and AI data tools as a national standard.

There is also agreement that greater focus on the use of data is not solely a technical skillset but a transformative framework that connects data and engineering practice with core aspects of professionalism, including ethics, inclusion and societal responsibility. Getting this right will position the UK as a leader in data-driven, sustainable engineering innovation. There is therefore a need for systematic alignment of DCE in engineering degree programmes with the Engineering Council's Accreditation of Higher Education Programmes (AHEP) framework.

The Royal Academy of Engineering, Alan Turing Institute and Lloyd's Register Foundation will continue to support the integration of DCE across tertiary education. We will repeat this exercise for engineering (and related) technician qualifications in further education, and we will continue to support this work in higher education through the following recommendations.

Recommendations

A summary of the recommendations is provided below.

1

Universities should adopt a two-tier Data-Centric Engineering outcomes map (baseline for all; advanced for some) with explicit assessment blueprints tied to application and critique

This study identifies a core Data-Centric Engineering baseline of knowledge, skills and behaviours for all engineering undergraduates regardless of discipline. This includes learning outcomes covering foundational statistics, data handling, analytics, and governance and ethics. We also identify a second tier of specialist DCE components, including advanced statistics, machine learning and AI, and large-scale data engineering, that should be offered as elective modules for those who want to specialise in this area.

2

Ensure integrated, practice-based DCE learning through laboratories and project-based approaches

To ensure DCE is included as an integrative component of undergraduate education, some creative thinking around existing laboratories and design spaces could open up opportunities to embed sensing, data management (transfer, conversion, pipelines, cleansing, etc.), uncertainty analysis, modelling, interpretation, and decision-making into existing laboratory, design, and capstone activities. Simulation and modelling tools will also play an increasingly important role, particularly for complex engineering products and systems. The effective use of laboratory spaces and digital software should be supported by tool-agnostic teaching packs and reusable project templates.

Where capacity is constrained, institutions should focus on a few well-designed laboratories or project experiences, plan technician and support workloads realistically, and combine hands-on work with validated simulations where necessary.

3

Build institutional capacity at scale: UK engineering higher education should build a community of practice to support teaching and learning in data engineering, including sector-wide micro-credentials on statistics for engineers, machine learning literacy, and the general growing use of AI in engineering education

A community of practice should be established to support the development and dissemination of professional development opportunities for academics and teaching staff – in statistics and data analytics for engineers, data engineering, ethics and governance around data and machine learning and AI literacy. The community of practice will help to support new pedagogy, spread good practice and reduce duplication of effort. Programmes such as the Academy's Visiting Professors can help support with real-life applications of data engineering from industry.

While not covered in this report, the Community of Practice could also provide a space for continued discussions around the growing use of generative AI and large language models (LLMs) more generally in engineering education (teaching and assessment).



4

Accreditation and standards alignment: propose explicit data engineering language in Engineering Council accreditation (AHEP 5) and other skills standards, using national exemplars to guide programmes

Professional bodies and sector partners should explore explicit data engineering language in the next iteration of the Engineering Council accreditation standard for higher education, AHEP 5 for Bachelor's (level 6) and Integrated (MEng) and standalone (MSc) Master's (level 7), supported by national exemplars and constructive alignment between outcomes, teaching activities, and assessment evidence.

Where possible, the learning outcomes included in this report should also be aligned with the recently produced UK Standard Skills Classification.

5

Ensure inclusion and widen participation in DCE pathways

Programmes should use diagnostics and just-in-time student support in subject areas such as maths, coding, and data handling, alongside inclusive communication and societally-relevant examples linked to sustainability, infrastructure, and safety-critical systems to broaden participation.

6

Enable sector coordination and a targeted implementation pathway

Use a maturity pathway model: Emerging → Adopters → Innovator. Using this model to support a staged approach aligned to the institutional typologies identified in the analysis:

Type A Emerging: enable minimum viable DCE (core outcomes + starter packs)

Type B Adopters: deepen practice-based learning (laboratories/projects)

Type C Innovators: scale exemplars and mentor others

7

Update teaching, learning and assessment to reflect DCE practice and use of artificial intelligence tools

As DCE and AI tools become more ubiquitous in engineering, universities need to redesign assessment to foreground judgement, explanation and critique. The Community of Practice should work collaboratively with the HE community to develop guidance on AI tools in teaching, learning and assessment.



1. Introduction

1.1 Background

In 2025, the National Engineering Policy Centre, led by the Royal Academy of Engineering, published its *Engineers 2030* report. It was a manifesto for setting out the future engineering skills needed for the twin challenges of meeting the UK's climate targets while continuing to grow the economy through innovation in technology-driven sectors.

The report, developed in consultation with a wide range of stakeholders across the whole of UK engineering, set out a vision and six principles on which the education and training of engineers and technicians should be based, shown in Figure 1.

Among the principles is the need for a more data- and digitally-fluent engineering workforce. The report states that engineers and technicians should be confident in understanding and using data and digital tools, as well as making ethical decisions about the use of emerging technologies. This detailed study builds on that principle and provides the foundation for ensuring future engineers will have the right skills to collect, use and apply data to improve engineering outcomes.

Engineering is a data-enabled discipline. From early recordings of information with a pencil and graph paper to 21st-century three-dimensional

modelling with computational fluid dynamics software packages, capturing and using data has always been a critical feature of engineering. But over the past two decades, we have seen significant advances in sensing, monitoring, computing power and analytics tools, including AI, leading to greater use of large-scale datasets across all aspects of the engineering lifecycle. Moreover, engineering systems have continually become increasingly complex, and our ability to understand physical phenomena across multiple scales and phases has led to a new paradigm.

Data-centric engineering (DCE) is thus growing as a field of engineering, focused on the integral use of data for engineering optimisation. It is an approach that fuses scientific and engineering principles with data as a fundamental component used for understanding engineering phenomena in processes, products and systems. It involves the collection, analysis and utilisation of data in unprecedented ways to inform decision-making; material characterisation; product and process design and optimisation; and management of large-scale, complex engineering systems.

Benefits are seen across the value chain, from innovation and new product development using simulation, modelling and design optimisation to manufacturing process control, condition

Figure 1: *Engineers 2030* vision and principles

Engineers 2030 Vision

By 2030, engineers play an urgent and pivotal role in sustainable growth, technological development and environmental regeneration with all sectors of engineering working inclusively and across fields.

Engineers are demonstrating leadership, creativity and technical excellence by implementing solutions that shape the future and enable society to navigate immediate challenges.

Engineers 2030 principles

- 1. Resilient and future-facing:** We navigate the changes that occur rapidly in our careers by embracing adaptability, continually developing our skills and knowledge, and collaborating across engineering disciplines.
- 2. Socially responsible and inclusive:** We draw on broad-ranging perspectives and

communicate widely, including with marginalised groups, to create, design and implement solutions that work for everyone.

- 3. Trusted by the public:** We recognise our professional ethical responsibilities in designing, creating and building a better future for people and the planet.
- 4. Integrated approach:** We manage and understand uncertainty in all its forms and work collaboratively to find creative and integrated solutions.
- 5. Data- and digitally-fluent:** We embrace digitisation, including artificial intelligence, and are skilled in working at the interface between the digital and physical worlds as they continue to merge.
- 6. Commercially and economically literate:** We generate knowledge within enterprise by using our technical knowledge and skills in creative ways for sustainable and equitable growth.

monitoring and predictive maintenance supporting the safety, reliability of plant and critical infrastructure, and a wide range of other benefits including supply chain management, minimisation of environmental impact, improvements in energy efficiency, and the management and control of complex systems.

As engineered systems become increasingly complex, their safety, reliability and efficiency require more knowledge than that which is available from underpinning classical theory. We are seeing increasing use of DCE, fusing fundamental physical and chemical laws with data-enabled and empirically-derived laws to manage and control these complex, nonlinear engineering systems.

To ensure the successful implementation of DCE across the wide range of engineering sectors, engineers and technicians need the knowledge and skills to use and analyse data more effectively. As such, there needs to be greater emphasis on the effective management, manipulation and analysis of data, and the use of statistical and optimisation tools, in particular AI, across engineering education at all levels.

The Alan Turing Institute, the UK's national institute for data science and artificial intelligence, has been championing DCE across science, engineering and technology. The work of the Turing Institute on DCE has been enabled through generous support from the Lloyd's Register Foundation,¹ an independent global safety charity that supports research, innovation and education to make the world a safer place. The Royal Academy supports this focus on DCE and is working in partnership with the Alan Turing Institute and Lloyd's Register Foundation to ensure that the UK can be a global leader in this critical development of engineering.

This report presents the case for greater embedding of DCE into university undergraduate engineering degree programmes across all disciplines. Future work will examine the skills needed for the technician workforce.

1.2 Purpose and scope

This work has set out to define the knowledge and skills required for effective use of DCE in industrial applications and promote these in engineering higher education. The aim is to define, assess and ultimately scale the integration of DCE into traditional, discipline-focused engineering degrees, in alignment with industry needs and emerging engineering practice.

Engineers increasingly need a new set of knowledge and skills in data to support the growth of data-driven decision-making across various engineering fields. These include:

- data acquisition
- handling and management
- data analytics and interpretation
- deeper understanding of probability and statistics
- programming skills for statistics
- the use of ML and AI tools for interrogation of large datasets
- ability to use a broad range of visualisation tools
- an understanding of how best to use data for decision-making and performance monitoring
- consideration of ethics, inclusion and privacy around the use of data.

The demand for DCE skills has also been strongly signalled by industry and national policy initiatives. Lloyd's Register Foundation's *Foresight review of big data*² helped start a major partnership with the Alan Turing Institute leading to the development of the DCE Programme, which has since engaged with leading universities and more than 30 industrial partners including NATS, Rolls-Royce and the Advanced Manufacturing Research Centre (AMRC), spanning sectors including aerospace, manufacturing, infrastructure, and transport. This demand is further reflected in the UK *National Data Strategy* and *The UK's Modern Industrial Strategy*, both of which emphasise the importance of strengthening data capabilities across engineering and manufacturing to support productivity, innovation and digital transformation.

The study recognises that DCE is not one single discipline, but a cross-cutting set of skills and approaches that are increasingly fundamental to contemporary engineering practice. In this context, the scope of the report is threefold.

- First, it maps the current curricular landscape by analysing programme and module documentation across a sample of UK universities.
- Second, it explores institutional perspectives and pedagogical practices through surveys and semi-structured interviews with academic and professional stakeholders.
- Third, it identifies strategic opportunities and gaps that can inform curriculum development, accreditation frameworks and national skills policy.

While the primary focus is on engineering disciplines, the report also considers interdisciplinary interfaces, recognising that DCE often draws from computing, data science and applied mathematics.

The findings are intended to support both policy and practice, guiding universities, professional bodies, the wider engineering profession and government stakeholders in developing a more coherent and forward-looking approach to embedding DCE within engineering education.

1 www.lrfoundation.org.uk

2 www.lrfoundation.org.uk/sites/default/files/2024-08/Foresight%20Review%20of%20Big%20Data.pdf

1.3 Project stages

This project has been developed to follow four phases:

Phase I: Identify the DCE knowledge and skills to teach in university degree programmes

This report presents the first phase, seeking to understand, at a granular level, the knowledge, skills and behaviours required for effective implementation of DCE in industry. This critical stage provides the foundation on which to build support for educators to develop effective future DCE curriculum and assessment.

Phase II: Assess the current provision of data engineering and AI data tools in Higher Education

This second phase attempts to undertake a gap analysis between the expected DCE knowledge and skills identified in Phase I and the current provision of DCE education in universities' undergraduate programmes of study.

Phase III: Build a community of practice among early adopters

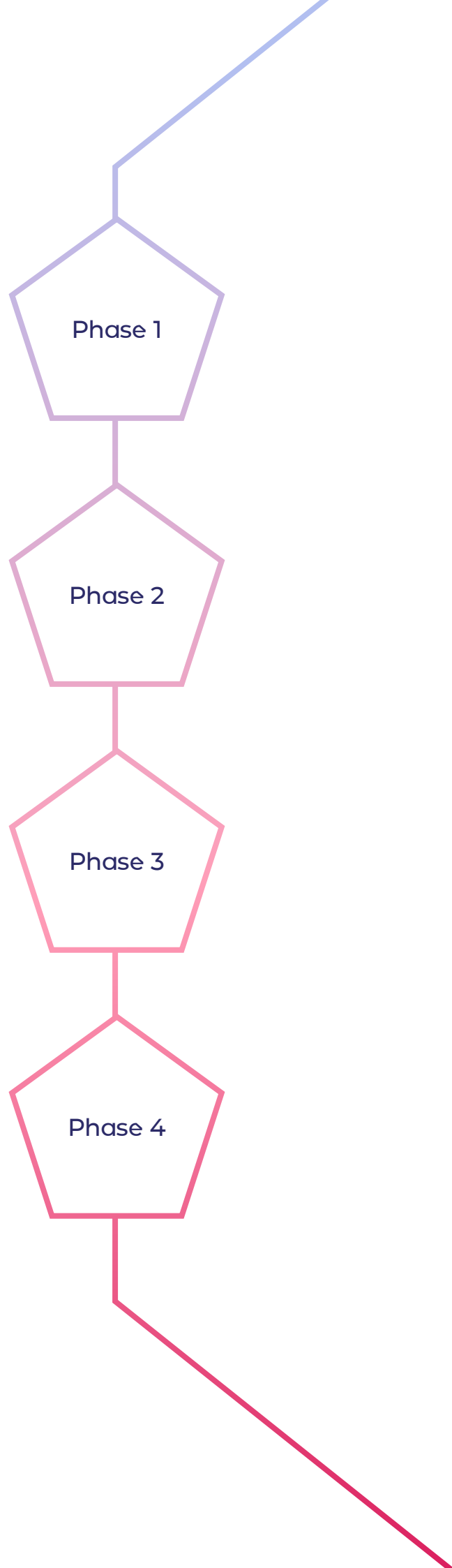
This phase, which will follow the publication of this report, is intended to create a network of universities currently working at the vanguard of DCE and create collaborative approaches to developing and disseminating good-practice teaching and learning of DCE across higher education to bridge the gaps identified in Phase II. This will include development of resources, support materials, CPD opportunities, pilots to test learning approaches, and so on.

We will also create the case for DCE as a core component in the next iteration of the Engineering Council Accreditation of Higher Education Programmes (AHEP).

Phase IV: Scale and influence national delivery

The final phase of the project will build on the early-adopter base to disseminate resources, teaching and learning approaches and other support to the wider HE community, through the Royal Academy of Engineering, Alan Turing Institute and wider engineering profession.

We will repeat the relevant elements of this methodology for technician qualifications in further education.



2. Methodology

As a preliminary step, an expert steering group was formed to advise on the content and direction of this study. Then, embedding DCE in UK higher education was considered as a process involving both curricular structures (for example, how topics are designed and delivered) and institutional strategies (for example, why and under what conditions they are adopted). To capture this dual dimension, we adopted a methodological framework consisting of a mixed-methods design, combining documentary analysis, survey research and semi-structured interviews^{3,4}. This approach enables us to:

- map the landscape of current DCE integration (quantitative/documentary)
- understand motivations, barriers and innovations (qualitative)
- build an evidence base for policy and curriculum design.

We followed a convergent parallel design by collecting qualitative (interviews) and quantitative (survey) data concurrently, analysed them separately and then integrate the findings for conclusions.

Expert steering group

The expert steering group's work was key in all the stages of this initiative. The steering group comprised experts and users of data engineering across different engineering disciplines and from academia and industry.

Steering group workshops support knowledge co-production and policy relevance, ensuring findings are actionable and aligned with HEIs' priorities^{5,6}.

Three workshops took place and aimed to:

- identify Learning outcomes across subject areas
- validate the identified Learning outcomes
- validate findings and co-construct practical recommendations in Phases III, IV.

Participants:

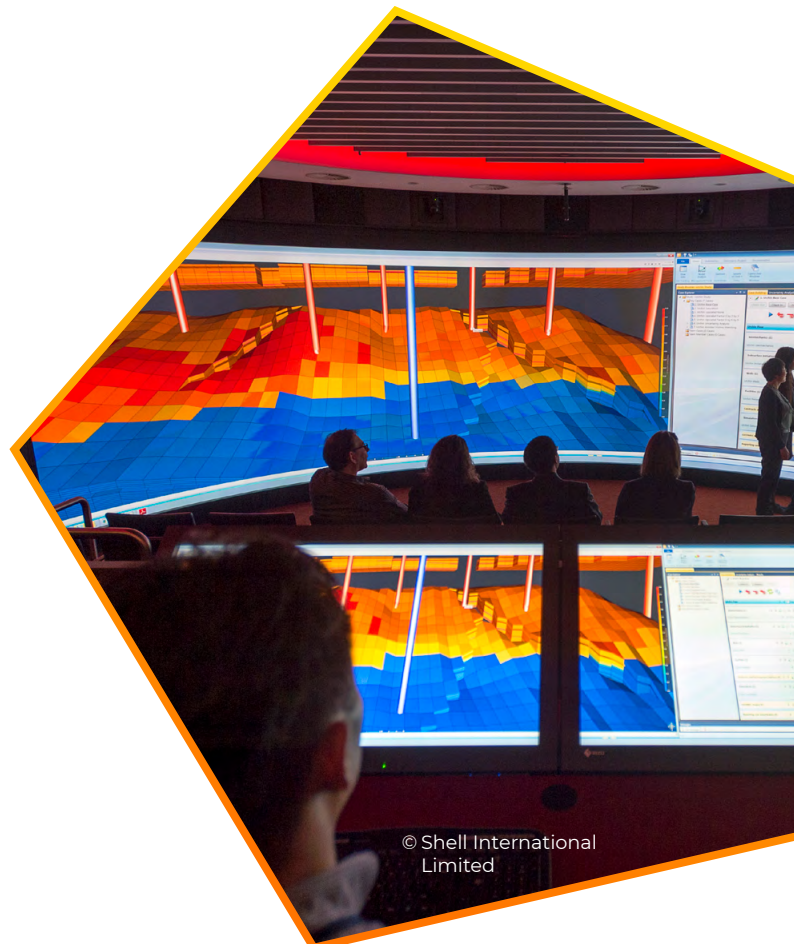
- academic leaders
- professional bodies
- industry representatives.

Method:

- structured facilitation techniques (for example, nominal group technique)
- present preliminary findings from survey and interviews
- identify gaps, priority areas and recommendations for scaling DCE integration.

Further detail on the methodology around data collection and analysis are provided in Annex 1.

- 3 Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th edn). Thousand Oaks, CA: Sage.
- 4 Creswell, J.W. & Plano Clark V.L. (2017). *Designing and Conducting Mixed Methods Research* (3rd edn). Thousand Oaks, CA: SAGE Publications.
- 5 Cornwall, A. & Jewkes, R. (1995) What Is Participatory Research? *Social Science and Medicine*, 41, 1667-1676. [http://dx.doi.org/10.1016/0277-9536\(95\)00127-S](http://dx.doi.org/10.1016/0277-9536(95)00127-S)
- 6 Bryson, J. M., Quick, K. S., Slotterback, C. S., & Crosby, B. C. (2013). *Designing Public Participation Processes*. *Public Administration Review*, 73, 23-34. <https://doi.org/10.1111/j.1540-6210.2012.02678.x>



© Shell International Limited

3. Data-centric engineering learning outcomes

This section presents a set of proposed learning outcomes for DCE within engineering undergraduate degree programmes.

These outcomes are intended to articulate the core knowledge, skills and practices that future engineers will require to operate effectively in increasingly data-driven engineering settings.

The learning outcomes for DCE were developed around a framework of the expected knowledge, skills and behaviours for undergraduate engineering students. They result from consultation with the expert steering and were further informed by the project's mixed-methods evidence base, including curriculum analysis, survey data and stakeholder interviews.

The proposed learning outcomes are organised across seven thematic areas:

- Foundational statistics
- Advanced statistics
- Data engineering and management
- Data analytics and visualisation
- Data governance, ethics and communication

- ML and AI for data analytics
- Additional AI skills for engineering

Together, these domains capture the technical and socio-technical competencies required to design, analyse and manage modern engineering systems that increasingly integrate physical models with data-driven methods.

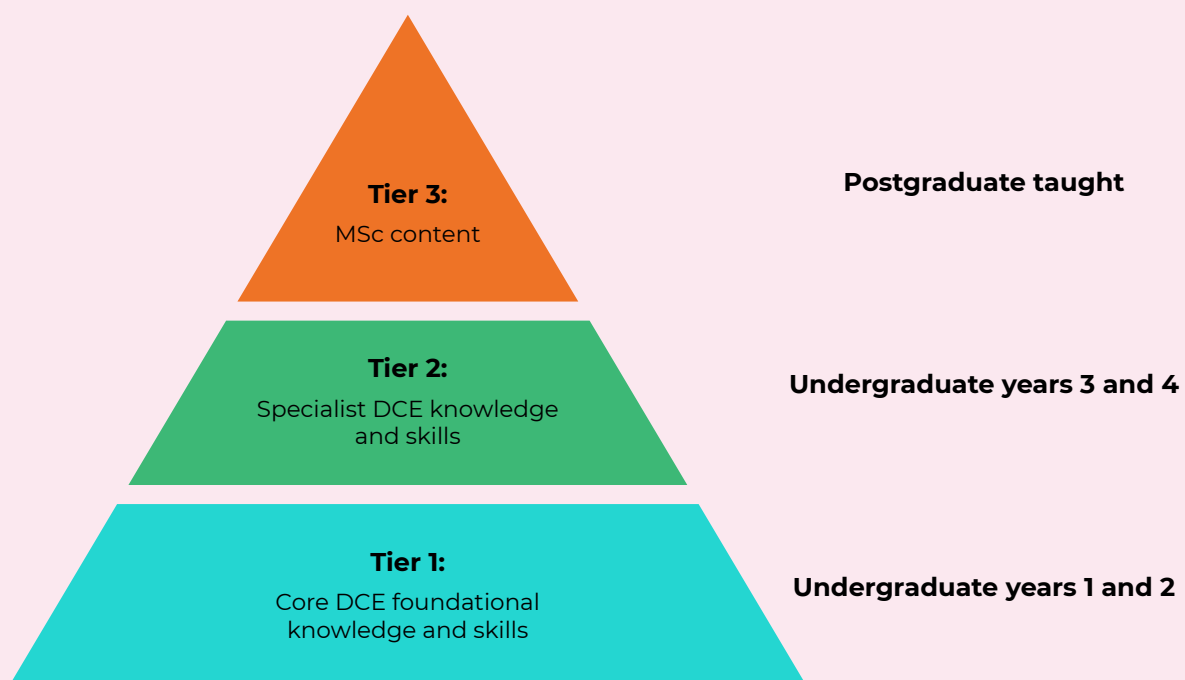
3.1 Tiered learning structure

The first stage of the framework creates a tiered structure, grouping learning outcomes into a 'core or foundational' tier, a specialist DCE tier and a final tier that could be used for creating stand-alone postgraduate taught programmes (if required).

Figure 2 presents the tiered structure.

Tier 1 identifies the foundational knowledge and skills that all engineers should have for working in new data-driven environments. While we do not specify when this content should be taught in an engineering degree, we would assume that it would be introduced as part of the core engineering curriculum, typically in Years 1 and 2 of undergraduate engineering degrees.

Figure 2: Tiered structure of the Data-centric engineering learning outcomes



Tier 2 includes additional or more specialist DCE content that would be more readily associated with some disciplines over others. For example, manufacturing engineering or chemical engineering programmes are more likely to have modules using statistical techniques and real-time data analytics for process control. This tier might also include modules that students choose as (electives) in the latter part of their degrees – Year 3 (and Year 4 in the case of MEng) in undergraduate degrees. These modules would enable those students to develop specialist skills for industry. Some of this Tier 2 content might also be learned or used by students individually or in small groups as part of final year design or capstone projects.

Tier 3 can be seen as the development of content for a stand-alone postgraduate taught MSc programme. It is likely to repeat the Tier 1 and 2 content for graduate students who have no experience of DCE, but may also stretch specialist learning further.

Tier 1 content is seen as the critical knowledge and skills that all future engineering graduates should know and use. In a later part of this project, we will engage the professional engineering community and the Engineering Council to promote inclusion of the core content in the next iteration of its Accreditation of Higher Education Programmes (AHEP 5). Tier 2 is likely to be more optional for individual engineering departments and higher education institutions to choose for their programmes. As we develop Phases III and IV of the project, we will build a community of practice around these more advanced elements of DCE and encourage more higher education institutions to adopt these learning outcomes in their undergraduate programmes.

3.2 ‘Modified’ Bloom’s taxonomy

The second element of the framework applies a modified Bloom’s taxonomy to the learning outcomes.

Each learning outcome is characterised by the level of cognitive engagement, ranging from:

- *Know*: Students have knowledge of the subject matter.
- *Apply*: Students can apply the subject matter to an engineering problem.
- *Understand*: Students have a theoretical understanding of the subject matter, including analysis and evaluation of the use of the subject matter.

Note: the original Bloom taxonomy reverses 2 and 3 above. However, for engineering, the application of statistical tools and techniques does not require a full conceptual understanding derived from first principles⁷.

This approach ensures that the proposed outcomes are not simply lists of topics but are framed in terms of progressive competence development, enabling alignment with curriculum design, teaching approaches and assessment strategies. Importantly, the identification of a core Tier 1 set of outcomes also provides a foundation for future dialogue with the engineering profession and the Engineering Council regarding the potential integration of DCE competencies into future iterations of the AHEP.

The tables in the section that follows present a structured articulation of the proposed DCE learning outcomes, organised by thematic area and level of competence, and are intended to provide a practical reference for universities seeking to embed DCE more systematically within engineering education.

3.3 Proposed DCE learning outcomes

The following tables present the proposed learning outcomes for DCE, structured across the competency areas identified through the project’s expert consultation. Together, **Tables 1–7** present the knowledge, skills and practices that engineering students should develop to operate effectively in data-driven engineering environments.

The tables progress from foundational statistical literacy (**Table 1**) to advanced statistical methods (**Table 2**), followed by competencies in data engineering and management (**Table 3**) and data analytics and visualisation (**Table 4**). The learning outcomes also incorporate essential governance, ethics, and communication capabilities (**Table 5**), before addressing the growing importance of machine learning and artificial intelligence (**Table 6**) and related advanced AI applications and systems integration (**Table 7**).

Each table specifies the expected level of cognitive engagement using the modified Bloom’s taxonomy adopted in this report and distinguishes between core outcomes expected of all engineering graduates and specialist capabilities that may be developed through electives, projects or postgraduate study.

⁷ Lucas, C. (2024) *A create-first pedagogy for Engineering*. UNSW (Sydney) festival of education, Nov 2024.

Proposed DCE Learning Outcomes

1



Foundational
statistics

2



Advanced
Statistics

3



Data
engineering and
management

4



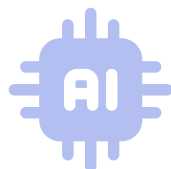
Data analytics
and
visualisation

5



Data governance,
ethics and
communication

6



ML and AI for
data analytics

7



Additional
AI skills



1. Foundational statistics

Table 1. Foundational statistical competencies for DCE education. These learning outcomes (descriptive statistics, probability theory, inferential statistics, and introductory regression analysis) cover essential elements required for engineers to interpret data, quantify uncertainty and draw evidence-based conclusions in engineering contexts. Each outcome is associated with a level of engagement using the modified Bloom's taxonomy ((1 = know, 2 = apply, 3 = understand)

Subtopic	Knowledge / Skills / Behaviour statement	Core	Specialist/ elective
Descriptive statistics	Calculations of measures of central tendency: calculate mean, median and mode	Understand	
Descriptive statistics	Calculations of measures of dispersion: range, variance, standard deviation, and interquartile range	Understand	
Descriptive statistics	Analysis of data sets through the creation of charts, graphs and summary statistics	Understand	
Probability theory	Concepts of probability, including conditional probability, independence and Bayes' theorem	Understand	
Probability theory	Various probability distributions (e.g., normal, binomial, Poisson) and their applications	Understand	
Inferential statistics	Different sampling techniques and the importance of sample size	Apply	
Inferential statistics	Point estimates and confidence intervals for population parameters		Apply
Inferential statistics	Hypothesis tests, including <i>t</i> -tests, chi-square tests, ANOVA, and understanding <i>p</i> -values and significance levels	Know	Apply
Regression analysis	Simple linear regression: Understand the relationship between two variables and the ability to use a simple linear regression model	Apply	Understand

Table 1 shows the proposed foundational statistical competencies that underpin DCE education and that are expected to form part of the core curriculum for all undergraduate engineering students. The learning outcomes cover essential elements of descriptive statistics, probability theory, inferential statistics and introductory regression analysis. Together, these competencies provide the statistical literacy required for engineers to interpret data, quantify uncertainty and draw evidence-based conclusions in engineering contexts. The table distinguishes between outcomes that should be achieved as part of the core undergraduate curriculum and those that may be further developed through specialist modules or electives. Each outcome is mapped to a level of the modified Bloom's taxonomy adopted in this report, indicating whether students are expected to know, understand or apply the concept within engineering problem-solving contexts. This foundational tier is intended to ensure that all engineering graduates possess the statistical foundation necessary to engage with more advanced DCE topics such as data analytics, machine learning and data-driven decision-making.





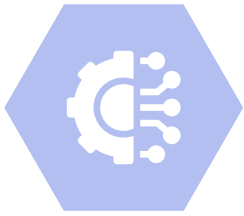
2. Advanced statistics

Table 2. Proposed learning outcomes for advanced statistical methods within DCE. Each outcome is associated with a level of engagement using the modified Bloom's taxonomy (1=know, 2=apply, 3=understand)

Subtopic	Knowledge / Skills / Behaviour statement	Core	Specialist/ elective
Regression analysis	Multiple regression: Understand and be able to extend linear regression to include multiple predictors	Know	Understand
Regression analysis	Nonlinear regression: Model and interpret nonlinear relationships	Know	Understand
Quality control and process improvement	Control charts (e.g., X-bar, R, p-charts) for monitoring process stability	Know	Apply
Quality control and process improvement	Process capability analysis: Assess process capability using C_{p^*} , C_{pk} indices		Apply
Quality control and process improvement	Six sigma methodologies and statistical tools for quality improvement		Know
Time series analysis	Identify and model trends in time series data	Know	Understand
Time series analysis	Seasonal decomposition: seasonal patterns and how to adjust for them	Know	Understand
Time series analysis	Forecasting techniques such as moving averages, exponential smoothing and ARIMA models		Apply
Design of experiments (DOE)	Principles of experimental design, including randomisation, replication and blocking	Know	Understand
Design of experiments (DOE)	Factorial design for experiments involving multiple factors	Know	Understand
Multivariate statistics	Principal component analysis (PCA): Ability to reduce dimensionality of data and identify key variables		Apply
Multivariate statistics	Cluster analysis: Grouping data points into clusters based on similarity		Apply
Multivariate statistics	Discriminant analysis: Techniques for classifying data into predefined categories		Apply
Nonparametric methods	Nonparametric Tests: Nonparametric statistical tests (such as Mann-Whitney U test, Kruskal-Wallis test and Wilcoxon signed-rank test)		Apply
Nonparametric methods	Bootstrap methods: Bootstrap techniques for estimating the distribution of a statistic		Apply
Statistical software proficiency	Statistical software: Mathematical or statistical software packages (such as R, MATLAB, SAS, SPSS and Python) (with libraries such as pandas, NumPy, SciPy) for performing statistical analysis	Apply	Apply
Programming and software skills	Programming Languages: Programming languages commonly used in engineering (such as Python, JavaScript, Java, MATLAB, C, and C++)	Apply	Understand

Table 2 presents the proposed learning outcomes for advanced statistical methods within DCE. These competencies extend the foundational statistical knowledge in Table 1 and focus on methods commonly required in data-driven engineering practice, including regression modelling, time series analysis, design of experiments, multivariate statistics, and statistical approaches to quality control and process optimisation. The table distinguishes between elements that may appear in the core engineering curriculum and those more typically delivered through specialist modules, electives or advanced project work. Each outcome is associated with a level in the modified Bloom's taxonomy. Collectively, these advanced competencies enable engineering students to analyse complex datasets, model system behaviour, design and interpret experiments, and apply statistical reasoning to optimisation and decision-making problems across engineering domains.





3. Data engineering and management

Table 3. Data engineering and data management. The focus is on engineering end-to-end data pipelines, database design and management. They also include the use of scalable technologies (for example, big data platforms and cloud computing) and the increasingly important cybersecurity. These learning outcomes equipped future engineers with competences to design and manage data infrastructure following reliable, secure and scalable practice.

Subtopic	Knowledge / Skills / Behaviour statement	Core	Specialist/ elective
Data pipelines	Data pipeline development: Data pipelines that collect, process and store data from various sources. This includes ETL (Extract, Transform, Load) processes	Know	Apply
Data collection	Data Acquisition: Methods to gather data from various sources, including sensors, databases and APIs	Know	Apply
Data cleaning	Data cleaning: Techniques to clean and preprocess data, handling missing values, outliers and ensuring data quality	Know	Understand
Database	Including data models, relational DB design; DB administration	Apply	Apply
Big data	Big data technologies: Big data tools and frameworks to handle and process large datasets efficiently	Know	Apply
Cloud computing	Cloud computing: Cloud platforms and their data services for scalable data storage and processing	Know	Apply
Cybersecurity	Practice of protecting systems, networks and data by means of technologies, processes and policies: cryptography, network & systems security, policies	Apply	Understand

Table 3 describes the proposed learning outcomes related to data engineering and data management, which form a critical infrastructure layer within the DCE framework. These learning outcomes focus on the processes required to acquire, organise, prepare, and secure engineering data so that it can be effectively analysed and used for decision-making. The learning outcomes cover key aspects of the engineering data lifecycle, including data acquisition from sensors and digital systems; development of data pipelines and ETL processes; data cleaning and quality assurance; database design and management; and the use of scalable technologies such as big data platforms and cloud computing. The table also emphasises the importance of cybersecurity and data protection practices, recognising that modern engineering systems increasingly operate within connected digital environments. As with other sections of the framework, the outcomes distinguish between core competencies expected of all engineering graduates and specialist capabilities that may be developed through advanced modules or project-based learning. Each outcome is associated with a level of cognitive engagement using the modified Bloom's taxonomy adopted in this report (1 = knowledge, 2 = application, 3 = theoretical understanding and evaluation). Together, these competencies ensure that engineers are equipped not only to analyse data, but also to design and manage the underlying systems that enable reliable, secure and scalable data-driven engineering practice.





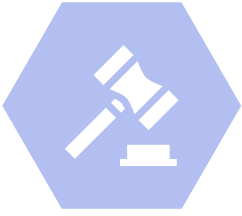
4. Data analytics and visualisation

Table 4. Data analytics, visualisation and domain-specific data skills learning outcomes. The outcomes focus on the practical competencies required for engineers to explore, interpret and communicate insights from data

Subtopic	Knowledge / Skills / Behaviour statement	Core	Specialist/ elective
Data analysis and interpretation	Exploratory data analysis (EDA): Initially explore datasets, visualise patterns and derive initial insights using packages/ tools such as Python (pandas, NumPy) or R	Know	Apply
Programming and software skills	Data analysis libraries: Libraries and frameworks such as pandas, NumPy, scikit-learn (for Python), or dplyr, ggplot2 (for R)	Know	Apply
Data visualisation	Visualisation tools to create clear and informative visualisations	Know	Understand
Data visualisation	Dashboard development: Interactive dashboards that allow stakeholders to interact with data in meaningful ways	Apply	Apply
Domain-specific data skills	Engineering-specific tools: Domain-specific data tools and software, such as finite element analysis, computational fluid dynamics and CAD software that integrates data analysis	Understand	Understand
Domain-specific data skills	Simulation and modelling: Creating and interpreting simulations and models relevant to their specific engineering field	Know	Apply

Table 4 shows the proposed learning outcomes related to data analysis, visualisation and the application of data-driven methods within engineering domains, forming a central component of the DCE competences. The outcomes focus on the practical competencies required for engineers to explore, interpret and communicate insights from data. Core elements include exploratory data analysis (EDA), the use of widely adopted data analysis libraries and programming environments, and the development of clear and informative visualisations. The table also highlights the ability to design interactive dashboards that enable stakeholders to interrogate engineering data in real time, reflecting the growing importance of data-driven decision-making in industry. Although the study recognises the role of domain-specific tools and simulation environments such as MATLAB, finite element analysis, computational fluid dynamics, and CAD systems, that increasingly integrate data analytics into engineering modelling and design workflows, it is equally important to emphasise that these tools and environment will be obsolete over time in a fast-changing computing landscape.





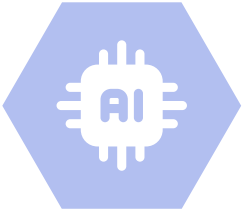
5. Data governance, ethics and communication

Table 5. Data governance, ethics and communication learning outcomes. These learning outcomes aim to equip learners with competences to understand the ethical, regulatory, and organisational implications of data use in DCE.

Subtopic	Knowledge / Skills / Behaviour statement	Core	Specialist/elective
Data ethics and privacy	Ethical considerations: Ethical considerations related to data collection, analysis and sharing	Apply	Apply
Regulatory compliance	Data privacy laws: Data privacy laws and regulations (e.g., GDPR, CCPA) to ensure compliance in handling data. Industry-specific data regulations	Know	Apply
Soft skills for data communication	Communication skills: Communicate (applying) data insights effectively to nontechnical stakeholders	Apply	Apply
Soft skills for data communication	Collaborative tools: Collaborative tools and platforms that enable teamwork on data projects, such as Git for version control	Apply	Apply
Data access and sharing	Trade-off collaboration: Compliance. Data access controls and role-based permissions.	Apply	Understand

Table 5 presents proposed learning outcomes related to data governance, ethical responsibility and communication within DCE. As engineering practice becomes increasingly data-driven, learners must not only analyse and manage data effectively but equally, as citizen and tomorrow’s engineers, understand the ethical, regulatory and organisational implications of its use. So, the learning outcomes address key competencies in data ethics and privacy, including responsible data collection, analysis and sharing, as well as awareness of relevant regulatory frameworks such as data protection legislation and industry-specific compliance requirements. The framework also emphasises the importance of effective communication of data-driven insights, ensuring that engineers can convey analytical results to diverse audiences, including nontechnical stakeholders involved in decision-making. In addition, the outcomes value the collaborative nature of modern engineering projects by including the use of collaborative tools and version-control systems that support teamwork in data-intensive environments. Finally, the table highlights the need to balance data accessibility with governance and security, through mechanisms such as role-based access controls and responsible data-sharing practices.





6. ML and AI for data analytics

Table 6 shows learning outcomes for ML and AI in DCE, covering core algorithms, model training and evaluation. It introduces deep learning architectures and GPU computing, and highlights the growing relevance of LLMs and generative AI for engineering analysis, research and data-informed decision-making.

Subtopic	Knowledge / Skills / Behaviour statement	Core	Specialist/elective
Machine learning (ML)	Algorithms: Various machine learning algorithms, including supervised learning (e.g., linear regression, decision trees, SVMs), unsupervised learning (e.g., clustering, dimensionality reduction), and reinforcement learning	Apply	Apply
Machine learning (ML)	Model training and evaluation: Training models on datasets, tuning hyperparameters and evaluating model performance using metrics such as accuracy, precision, recall, F1 score, and ROC-AUC	Know	Understand
Machine learning (ML)	Tools and frameworks: Machine learning libraries and frameworks such as scikit-learn, TensorFlow, Keras, and PyTorch for developing and deploying ML models	Apply	Understand
Deep learning	Neural network architectures: Different types of neural networks (for example convolutional neural networks for image processing, recurrent neural networks for time series and natural language processing) and their applications	Know	Understand
Deep learning	GPU computing: GPU acceleration for training deep learning models, using tools such as CUDA and cuDNN to optimise performance	Know	Apply
Large language models and generative AI	Advanced AI systems to understand, generate and manipulate human language. Research and writing tools	Apply	Understand

Table 6 outlines the proposed learning outcomes related to ML and artificial intelligence within the DCE. These competencies reflect the growing role of AI-driven methods in analysing complex engineering datasets, modelling system behaviour and supporting data-informed decision-making across the engineering lifecycle. The learning outcomes span aspects including an understanding of core ML algorithms (both supervised and unsupervised), the processes involved in model training and evaluation. The framework also introduces more advanced concepts associated with deep learning architectures, including neural networks and the computational infrastructure required to train them efficiently, such as GPU-based computing. In addition, the table acknowledges the growing integration of LLMs and generative AI (GenAI) tools within engineering practice, recognising their increasing role in engineering research, technical documentation, knowledge management, and decision-support workflows.





7. Additional AI skills

Table 7. Additional artificial intelligence and systems integration learning outcomes. This table outlines learning outcomes for advanced AI applications in engineering, including NLP, computer vision, robotics, and cloud-based AI deployment. It emphasises ethical AI practices and systems integration, highlighting how statistics, data engineering and AI interact in complex systems. Outcomes are classified as core or specialist and aligned with the modified Bloom's taxonomy.

Subtopic	Knowledge / Skills / Behaviour statement	Core	Specialist/ elective
Additional AI-related skills	Natural language processing (NLP): NLP techniques for text analysis, sentiment analysis and language modelling using tools such as NLTK, spaCy and Hugging Face Transformers	Know	Apply
Additional AI-related skills	Computer vision: Computer vision techniques for image and video analysis, using libraries such as OpenCV and frameworks such as TensorFlow and PyTorch	Know	Apply
Additional AI-related skills	Robotics and automation: Integration of AI in robotics for autonomous systems, using tools such as ROS (Robot Operating System) and simulation environments such as Gazebo	Know	Apply
Additional AI-related skills	AI ethics and fairness: Ethical considerations in AI, including bias mitigation, fairness, transparency, and explainability of AI models	Apply	Apply
Additional AI-related skills	Cloud computing for AI: Deploying AI solutions on cloud platforms such as AWS, Google Cloud and Azure, using their AI and ML services	Know	Apply
Systems integration and functional threads	A basic understanding of how the statistical analysis, data engineering and artificial intelligence all come together for an applied and useable system – how is the integration achieved and how do you qualify and trust the outputs	Apply	Apply

Table 7 presents additional learning outcomes related to advanced AI applications and the integration of data-driven technologies within engineering systems, extending the machine learning competencies proposed in the previous section. The outcomes highlight areas of AI that are increasingly relevant to engineering practice, including NLP, computer vision and the integration of AI within robotics and automated systems. These capabilities reflect the expanding role of intelligent systems in engineering design, monitoring and control across sectors such as manufacturing, infrastructure and autonomous technologies. The framework also emphasises ethical considerations in AI, including issues of bias, fairness, transparency, and explainability, recognising that responsible deployment of AI is an essential component of modern engineering practice. In addition, the table addresses the use of cloud-based AI platforms for scalable deployment of data-driven solutions, reflecting current industry practices in large-scale data processing and model deployment. Finally, the learning outcomes highlight the importance of systems integration, ensuring that students develop a conceptual understanding of how statistical analysis, data engineering and artificial intelligence interact within complex engineering systems, and how the reliability and trustworthiness of such systems can be evaluated. As with other elements of the framework, outcomes are classified according to whether they form part of the core knowledge expected of all engineering graduates or specialist capabilities that may be developed through advanced study.



Taken together, **Tables 1-7** define a structured framework of proposed DCE learning outcomes, covering statistical foundations, data management, analytics, governance, and artificial intelligence. This framework establishes the reference against which current provision across UK engineering programmes is examined in the following chapter through analysis of the collected quantitative and qualitative data.

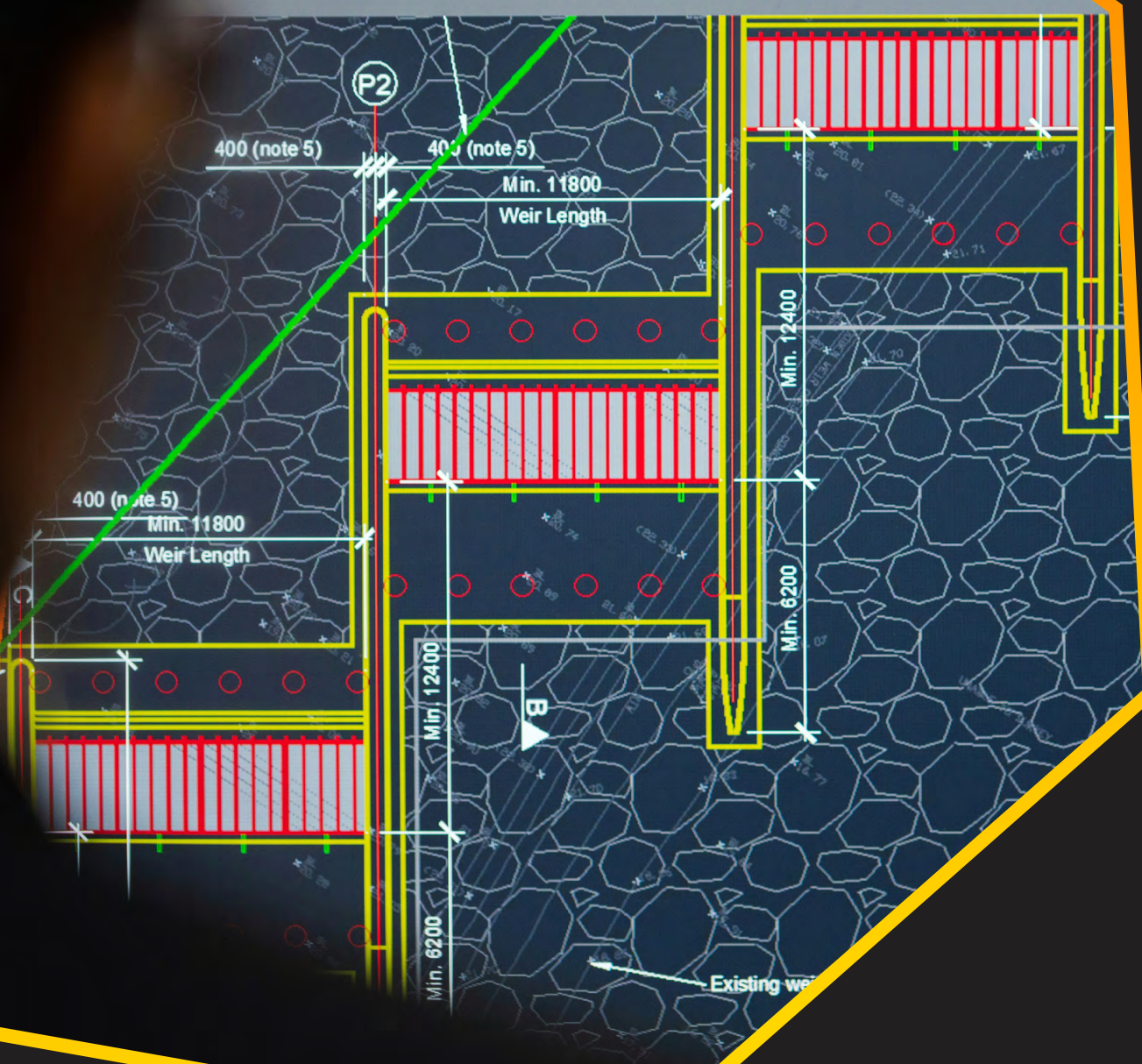
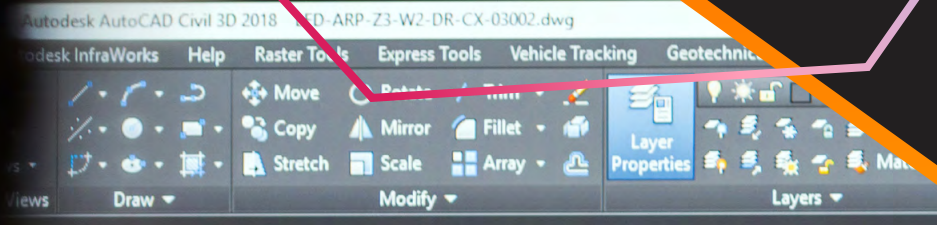
3.4 Obsolescence of the learning outcomes

The study recognises that the area of DCE is developing rapidly with AI, LLMs and agentic AI likely to have a significant impact on the relevance of the learning outcomes.

We have tried to generalise the learning outcomes and identify the key foundational knowledge and skills required regardless of advances in technologies used. For example, we have not included any specific softwares, data platforms or AI models as the pace of change of development means these may be obsolete in a relatively short time frame.

The other challenge is that some AI tools may make some of the learning outcomes redundant. For example, data cleaning is increasingly being undertaken by AI. However, it is nonetheless important that engineers understand that raw data from sensors and so on can have, for example, errors and/or missing data and that, regardless of technology platform, it needs to be cleaned to be of effective use.





4. HE capacity to deliver the learning outcomes

Building on the proposed DCE learning outcomes, this chapter examines the extent to which the teaching and/or learning of these competencies is currently reflected in UK engineering higher education. Drawing on the mixed-methods approach described in the methodology, this chapter presents the results of both quantitative analysis of survey data from 60 engineering departments and qualitative insights from stakeholder engagement. Together, these analyses provide an evidence-based view of current curricular provision, institutional capacity and emerging practices related to the integration of DCE across UK engineering higher education.

4.1 Quantitative analysis

4.1.1 Descriptive analysis

A nationwide survey was conducted to provide an overview of how the selected DCE topics (foundational statistics, advanced statistics, data management, data analytics, governance/ethics, and

AI/ML) are taught or learned across UK institutions.⁸ By doing so, we established a baseline map of DCE provision across disciplines (for example, mechanical, civil, electrical). The main findings of the descriptive analysis are presented in this section, with further details provided in Annex A.

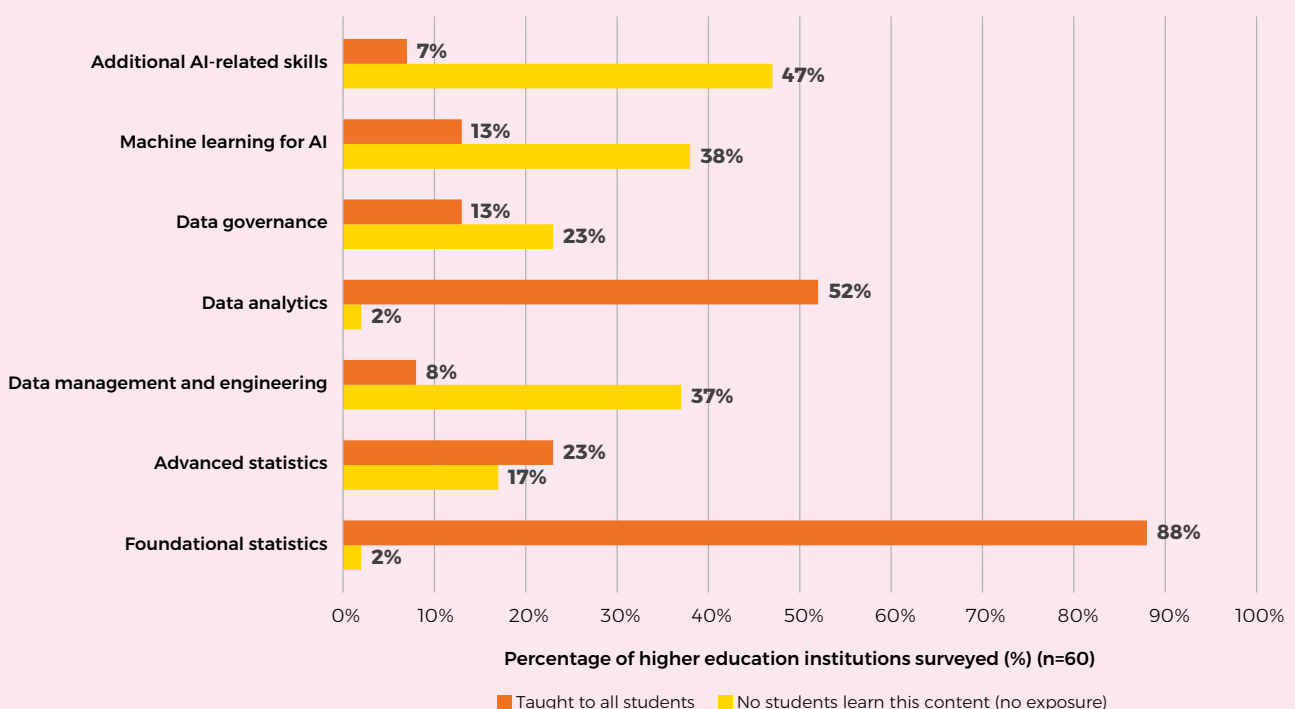
Figure 3 shows a landscape where data-handling literacy (foundational statistics, analytics and ethics) is reasonably well embedded across UK engineering, but for more advanced and infrastructural aspects of DCE (data engineering, use of AI), there is no teaching of them at all (no exposure).

This baseline map of DCE provision showed that:

- Foundational statistics and data analytics are widely embedded and often core (high average coverage, many reporting 'core for all students').
- Governance/ethics is present but uneven, with

8 It is important to note that learning can take place without teaching taking place, for example, in self-directed individual or group project work.

Figure 3: Extent of DCE coverage across engineering departments in higher education institutions HEIs



a worrying minority of departments reporting no coverage (overall, about half of students, on average).

- Advanced statistics, data management, ML, and advanced AI skills are present but often confined to electives, options or project work rather than forming part of the core. Many departments either offer these to small subgroups of students or not at all.

Figure 4 shows how the different elements of Data-Centric Engineering are delivered in higher education institutions when the thematic areas are provided to students.

Similarly, when looking at institutions' *capacity* versus *coverage*, data show the following three categories (shown in **Figure 5**):

- topics with high coverage and high capacity: Foundational statistics, data analytics
- topics with moderate coverage but mixed capacity: Governance/Ethics, ML
- topics with lower coverage and weaker capacity: Data management & engineering, additional AI-related skills.

Across disciplines, mechanical and electronic/electrical engineering departments appear to be the most advanced in terms of embedding ML and broader AI topics, while more specialised disciplines

often have pockets of activity rather than systemic coverage.

More generally, we found that the majority of universities are exploring how to use AI effectively in teaching and learning, but for assessment, a greater proportion of institutions are allowing students to use AI tools in assessment across a range of subjects, as shown in **Figure 6**.

Additionally, DCE inclusion does not systematically depend on discipline, but delivery depth and pedagogy do depend on institutional capacity. Strengthening staff skills and improving resources for applied data work will enable consistent DCE adoption across all branches of UK engineering.

If the UK's engineering ambitions aim to mainstream data-enabled engineering then investment may need to focus on:

- staff development
- supporting project infrastructure (datasets, computing resources, demonstrator systems)
- sharing exemplars from high-capacity departments to lower-capacity ones.

Because discipline effects are weak, national interventions can reasonably be framed at the level of 'engineering education as a whole' with some tailoring for local context rather than discipline-specific mandates.

Figure 4: Modes of teaching and learning of data engineering and AI in universities

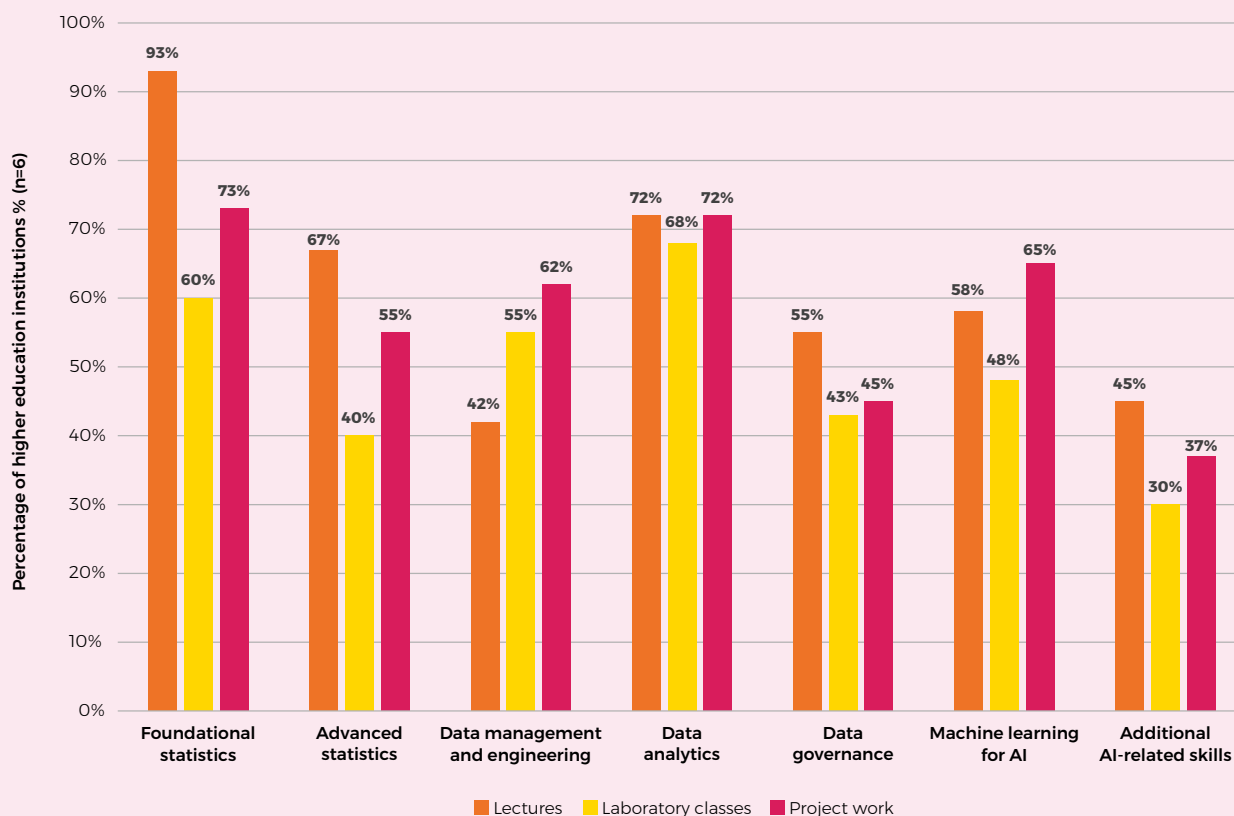


Figure 5: Capacity of higher education institutions to deliver various aspects of data engineering and AI



Figure 6: Use of AI tools in teaching, learning and by students in assessment



4.1.2 Multivariate analysis (exploratory factor and cluster analysis)

A multivariate analysis was conducted to identify institutional typologies in DCE provision. That means identifying the characteristics that shape DCE provision and group UK engineering departments into distinct institutional typologies. Details on methods and results are found in Annexes A and B.

DCE provision was identified and categorised in terms of three broad characteristics:

- *Pedagogical integration*: How pervasively, across topics and modes, DCE appears in the curriculum.
- *DCE/AI capacity*: How confident departments feel about their ability to teach these topics, particularly advanced statistics and ML and AI.
- *Practice-based integration*: The extent to which DCE is delivered via projects and laboratories rather than only via lectures.

These factors give a compact, interpretable representation of the survey data, which we can then use to group institutions. The cluster analysis surfaces three groups of institutions which contain a mix of integration levels in the simple ‘core topics’ sense, the structural differences residing more in how and with what capacity DCE is delivered.

Type A (Emerging / low-integration departments): These departments are traditional in their curriculum structures and appear to be at an early stage of DCE adoption. They have some capacity but relatively little DCE in the taught curriculum, especially for ML and project-based work.

Type B (Content-led DCE adopters): These departments have moved fastest to embed DCE into taught content, especially via lectures and structured modules. They are firmly ‘emerging adopters’: DCE is visible across the curriculum, but

their capacity and practice-based integration lag behind the most advanced institutions.

Type C (High-capacity, practice-based DCE innovators): These are the leading-edge institutions: they combine strong capacity with a clear emphasis on project-based and lab-based DCE learning. ML and advanced statistics are not just talked about; they are applied substantially in project work.

Proposal: Considering that the main fault line is not discipline, but how departments translate capacity into pedagogy, we would suggest:

- Helping Type A departments take first steps (e.g. foundational stats + basic analytics in core modules).
- Enabling Type B departments to shift from lecture-led to project-rich DCE.
- Showcasing the practice of Type C innovators (curriculum exemplars, staff development, resource models) as national reference cases.

4.1.3 Summary of quantitative analysis

This section synthesises all quantitative analyses in the survey of *Data-Centric Engineering (DCE) education in UK engineering higher education*.

Table 8 consolidates findings from the quantitative analysis. Details are presented in Annex B.

4.2 Qualitative: Thematic analysis

To gain deeper insights into how and why DCE topics are integrated or not within curricula, and in alignment with best practices in educational research and curriculum studies⁹,

9 Cohen, L., Manion, L., & Morrison, K. (2017). *Research Methods in Education*. New York, NY: Routledge. <https://doi.org/10.4324/9781315456539>

Table 8: Summary of findings from the quantitative analysis of HE engineering departments’ readiness for teaching DCE

Theme	Strategic Implication
1. UK DCE education shows strong foundations but limited depth.	Build beyond basics: national focus should shift from <i>literacy to applied data capability</i> .
2. Institutional capacity, not discipline, drives DCE innovation.	Policy interventions should target capacity building rather than discipline-specific mandates.
3. Project-based learning is the hallmark of advanced DCE maturity.	Incentivise practice-based curricula through funding and exemplars.
4. Three institutional archetypes exist.	Support pathways: (A) Introductory adoption, (B) Pedagogical enhancement, (C) Showcase exemplars.
5. Cross-method coherence strengthens reliability.	Findings are robust, methodological convergence supports confidence in the baseline map.

we conducted semi-structured interviews with a range of senior leaders in university engineering departments across different higher education mission groups. We spoke to academics designing and delivering relevant modules, curriculum leads and senior teaching staff. This was followed by a thematic analysis of the data collected. These interviews provided a rich, contextualised understanding that explains patterns found in the survey. The following sections present the findings.

4.2.1 Commonalities

1. Pedagogical integration: embedding DCE, not adding it

A dominant theme across all institutions is that interviewees reject a stand-alone, add-on model. There is a shared commitment to *embedding* DCE principles within existing engineering curricula rather than creating stand-alone modules. DCE should be a thread that runs through laboratories, design projects and capstones; students encounter sensing, data acquisition, data cleaning, visualisation, and interpretation in discipline-authentic contexts. Interviewees prioritise course redesign around authentic tasks, not new modules, so DCE becomes part of routine engineering judgement. It reflects a pedagogical preference for integration that minimises curriculum overload, aligns with accreditation expectations and fosters more authentic learning experiences.

Proposal: Prioritise course redesign around existing tasks, not new modules, so DCE becomes part of routine engineering judgement.

2. Core v specialist competencies

A consensus model distinguishes foundational competencies for every degree student (data handling, elementary statistics, basic data engineering) from advanced competencies for specialists or upper-year electives (for example, regression families, optimisation, ML/AI); this is articulated as a progression across undergraduate years and into postgraduate provision. This is in line with the curriculum scaffolding approach ensuring progression from basic literacy to disciplinary mastery¹⁰.

Proposal: Adopt a two-tier outcomes framework, ensuring minimum DCE literacy for all while enabling specialisation through electives and projects.

3. Ethical, legal and societal dimensions of DCE

DCE is a natural vehicle to teach ethics, privacy,

GDPR, bias, and responsible AI. Across all HEIs institutions interviewees, ethics and governance were seen as central, not peripheral to DCE education. HEI-C argued for ethics to be embedded contextually “as professional reasoning”, rather than taught as an “add-on”, particularly around AI, privacy and responsible innovation. HEI-D treats ethical and professional considerations, including data privacy and accountability, as embedded learning outcomes from first year onwards. They embed ethical accountability from Year 1 and cautions against the “black-box” use of ML. HEI-A raised issues related to AI in assessment and academic integrity, stressing the need for institutional guidance. HEI-B also flagged inclusion and accessibility as ethical considerations in framing DCE content.

Proposal: Define assessable learning outcomes on ethical data practice and uncertainty quantification across the spiral curriculum.

4. Institutional capacity, staff development and cultural change

Interviewees consistently noted that staff confidence and institutional readiness are critical to successful DCE implementation. All university representatives interviewed reported variable staff confidence (especially in probability/statistics and ML) and highlighted the need for CPD and shared resources. HEI-C reported uneven levels of statistical literacy among faculty and suggested reframing content around the engineering language of ‘uncertainty and risk’ to encourage greater buy-in and build consensus. HEI-A pointed to capacity gaps in less data-intensive disciplines and emphasised the need for CPD resources and sector-wide support. HEI-B described DCE integration as a ‘cultural shift’, requiring faculty buy-in and sustained professional development. HEI-D noted the challenge of staying up to date with rapidly evolving software tools while focusing on fundamental skills. HEI-A and HEI-D stress uneven readiness across departments and the need for open CPD toolkits and practice communities. These views convey the idea that pedagogical innovation is inseparable from institutional change.

Proposal: Pair curriculum change with systematic CPD, exemplars and coaching; community of practice on assessment redesign in an AI-pervasive environment.

5. Accreditation and policy levers

Interviewees linked DCE adoption to the evolution of accreditation frameworks, particularly the AHEP.

HEI-C and HEI-A explicitly discussed the constraints and opportunities presented by AHEP 4, with optimism that AHEP 5 will allow greater emphasis on DCE. HEI-A proposes explicit data/AI references in outcomes (M1/M2 and C1/C2) and placement at FHEQ L6 to ensure universal proficiency. HEI-B saw

10 Biggs, J. B. & Tang C. (2003). *Teaching for quality learning at University* (5th edn). Houston: Open University Press.

alignment with accreditation as essential for scaling adoption nationally. They suggested the need for national guidance and a community of practice to influence AHEP 5 and normalise embedded DCE. HEI-D expressed willingness to engage with shaping AHEP 5 through early adoption and resource development. HEI-A noted that accreditation already privileges application (ethics, sustainability, simulation) over encyclopaedic science coverage, undermining “we can’t fit it in” arguments.

These considerations highlight the role of regulatory and policy frameworks as enablers of pedagogical change, a finding consistent with previous national curriculum reform efforts in engineering education.

Proposal: Aligning programme learning outcomes to an AHEP-consistent DCE spine, using accreditation language to legitimate rebalancing.

6. Inclusivity and accessible entry routes

HEI-B highlighted inclusive framing to avoid discouraging students with weaker coding backgrounds; language matters for participation of under-represented groups. HEI-F interviewee added that uneven prior maths/practical skills demands bridging toolkits and careful scaffolding.

Proposal: Provide on-ramp resources (diagnostic support, just-in-time maths/data ‘toolkits’) and inclusive task design from Year 1.

7. Resources, laboratories, and scale: embedding under real constraints

The HEI-F interview surfaced hard constraints, technician capacity, laboratory throughput (multiple repeats of lab exercises to provide for all students), equipment at scale, post-COVID drift to “dataset-only” laboratories and warned that authentic “hands-on learning” is expensive and yet critical to connect abstract mathematics with more concrete engineering applications.

Proposal: Prioritise high-leverage laboratories (fewer, richer), instrument existing experiments for data work, and support with virtualisation only where it complements but not replaces hands-on.

8. Software-agnostic fundamentals & assessment redesign

HEI-C urges software-agnostic teaching (principles over brand-specific tools) and elevates Bayesian reasoning/uncertainty. HEI-D echoes focus on transferable modelling and critical evaluation rather than tool-chasing. HEI-A and HEI-F emphasise assessment challenges in the LLM era and the need to evidence authentic engineering reasoning.

Proposal: Redesign assessments to foreground judgement, explanation and validation; permit tool use but assess understanding and decision-making.

9. Industry alignment & employability

HEI-E underscores employer demand for DCE skills and shows project-led mastery (for example, data cleaning for industrial rigs), while acknowledging uneven depth across student choices; staff capability often tracks research strengths. HEI-B links DCE to employability narratives and recommends national showcases and early-adopter pathways.

Proposal: Expand industry-anchored projects and visiting-professor partnerships to keep tasks authentic and skill signals legible to employers.

4.2.2 Points of divergence (productive tensions)

Across interviewees the emerging commonality from a strategic view is that DCE should be framed as a transformative lens for engineering practice. That is, a mechanism for engineers to think and decide under conditions of uncertainty, rather than an additional topic to add in the curriculum. HEI-C’s forthcoming *Design and Data-Centric Engineering* core paper epitomises this reframing. The HEI-D redesign similarly integrates data/science/ethics from Year 1. DCE is treated as a capability architecture that re-organises existing teaching, enabling simplification (not accretion) of curriculums.

However, diverging views have also surfaced:

- For example, where to locate statistics/ML depth? Some favour widespread light-touch exposure; others prefer concentrated advanced options.
- Space v substitution: Some universities argued for taking content out to make room for DCE, highlighting its increasing importance in engineering practice; others prefer embedding within existing laboratories with minimal timetable change.
- Facilities realism: HEI-F cautions that scaling authentic laboratories is hard without technicians and investment in capital equipment; this may temper ambitious embedding plans.



5. Discussion and recommendations

5.1 Discussion

Data engineering is now a fundamental requirement for engineering practice in the twenty-first century. Added to that, the amount of data that will be generated by products, processes and systems will require ever more sophisticated tools to be used by engineers. This report has set out what we believe are the knowledge, skills and behaviours required by all engineers to apply data engineering effectively in industry.

The proposed DCE learning outcomes translate the main goal of embedding DCE as a cross-cutting capability, into a concrete, assessable curriculum framework that universities and accreditors can act on. By presenting a tiered progression (foundational core for all students, specialist/elective content in later years and extension into postgraduate provision) and mapping outcomes across key domains (statistics, data engineering/management, analytics and visualisation, governance/ethics, and AI/ML), the outcomes provide a national reference point for curriculum design, benchmarking and gap analysis. And importantly, the outcomes are framed in terms of knowledge, application and deeper understanding (via a modified Bloom's taxonomy), which supports constructive alignment between teaching, assessment and programme aims which is an essential requirement for a robust quality assurance.

We have highlighted in this work that technology obsolescence will be a problem and so we have tried as much as possible to draw out the fundamentals of data engineering, but we have identified current AI and big data tools as exemplars of current practice to aid readers.

We have also been keen to highlight the importance of data governance and ethics in this work. Ethics is an important element of an engineer's education. It is a requirement in the Engineering Council's current accreditation guidelines for undergraduate degrees. There have been an increasing number of questionable ethical practices across engineering in recent years, including around the software developments for the UK Post Office. Data engineering, and all the aspects of GDPR are an excellent vehicle for teaching ethics and should be embraced by engineering departments.

The analysis of UK higher education to deliver against these learning outcomes is overall positive. All the universities that we surveyed are delivering core, foundational statistics – the mathematics of data. Indeed, most students will have covered

much of this learning before even arriving at university.

But as we dig deeper into the requirements of DCE, we start to see a mixed picture – both from a disciplinary and institutional perspective. As we might expect, subjects such as chemical engineering and manufacturing and production engineering will have a greater bias towards statistical techniques, as tolerances, process optimisation, quality control, and so on are all important learning outcomes for graduates entering those industries. Subjects such as civil engineering have historically not focused so heavily on these the statistics of process control because of the 'one-off' nature of the engineering products derived from the discipline. However, as we move towards a world where asset management through infrastructure sensing and condition monitoring becomes increasingly important, we will necessarily need to see a shift in the education of engineers working in the built environment to increasingly deal with data – as we have already seen with the introduction of Building Information Modelling, measures of sustainability, carbon emissions and so on. The research found that certain subjects lend themselves to greater use of software, programming, ML and AI. Electronic engineering notably standing out as a strong early adopter.

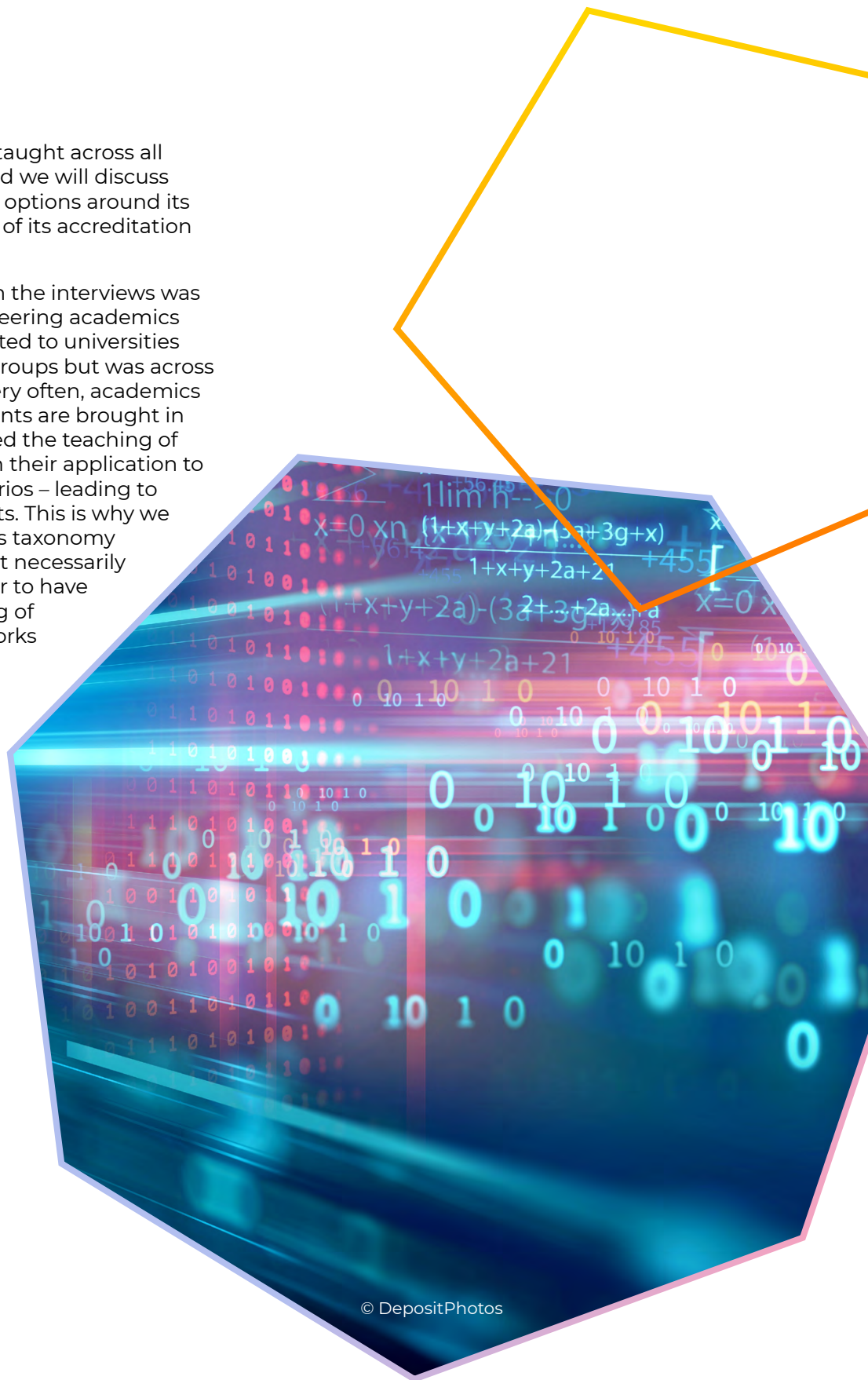
The reporting of the qualitative results sets out several proposals. These are presented in the recommendations below, but a number are picked out here for further elaboration.

There was a consistency of views from interviewees that engineering curriculums are already over-full and it is strongly preferred that data engineering is integrated into the curriculum rather than bolted-on as yet more teaching. Pedagogy will play an important part in achieving this. For example, in laboratory classes and project work, students learn by doing – we would encourage departments to allow students to set up experiments themselves, understanding data acquisition through the use of sensors and their positioning on equipment, capturing data, using A-to-D converters, storing and managing the data (perhaps in the cloud), removing noisy data to undertake analysis of clean data. We do of course recognise that time, resource and capacity all act against this ambition. Accreditation was brought up as a subject here. There were mixed views on the use of accreditation to drive DCE. While all recognised it is a powerful tool for compelling departments to adopt the learning outcomes, there was concern that departments are already stretched to deliver against current requirements. There is clearly

a need to ensure that DCE is taught across all engineering departments, and we will discuss with the Engineering Council options around its inclusion in the next iteration of its accreditation guidelines in AHEP 5.

A final point to highlight from the interviews was the lack of readiness of engineering academics to teach DCE. This wasn't limited to universities from any particular mission groups but was across the board. We learned that very often, academics from mathematics departments are brought in to teach statistics and favoured the teaching of the fundamentals rather than their application to real-world engineering scenarios – leading to disengagement from students. This is why we believe the 'modified' Bloom's taxonomy works for engineering. It is not necessarily a requirement for an engineer to have a fundamental understanding of why a statistical technique works from first principles before they can use it, but just to know how to apply it and when it is best to apply it. A useful comment on encouraging engineering staff to teach statistics was to frame it in the language of engineering: 'uncertainty and risk'. But regardless, universities across the board will need support in giving staff the confidence to teach DCE. This will be a key element of this project going forward.

The next section lays out the recommendations drawn from the research.



© DepositPhotos

5.2 Recommendations

The proposals are grouped by the aspect of provision they address (curriculum expectations, pedagogy and infrastructure, staff capacity, assessment, accreditation and standards, inclusion, and sector enablement) to support action.

1. Universities should adopt a two-tier Data-Centric Engineering outcomes map (baseline for all; advanced for some) with explicit assessment blueprints tied to application and critique

- 1.1 Adopt a two-tier DCE outcomes map (core + specialist). Establish a national expectation that all engineering graduates achieve a core DCE baseline (e.g., foundational statistics, data handling, analytics, governance/ethics), with specialist pathways (advanced statistics, ML/AI, large-scale data engineering) delivered through options, specialisms, and capstone projects.
- 1.2 Define “minimum viable DCE” at degree level, with exemplars. Publish short, discipline-agnostic guidance describing a minimum viable content for DCE at undergraduate level, including example module structures, assessment approaches and workload models so that departments can implement without redesigning entire programmes.

2. Ensure integrated, practice-based DCE learning through laboratories and project-based approaches

- 2.1 Apply creative thinking around existing laboratories and design spaces to open up opportunities to embed sensing, data management (transfer, conversion, pipelines, cleansing, etc.), uncertainty analysis, modelling, interpretation, and decision-making into existing laboratory, design, and capstone activities.
- 2.2 Increase the use of simulation and modelling tools, digital twins etc. to enable students to experience data analytics for complex engineering products and systems.
- 2.3 Expand high-quality project templates and tool-agnostic teaching packs. Develop reusable, modular “DCE project packs” that can be adopted across institutions (datasets, briefs, marking criteria, academic integrity guidance, accessibility adjustments). This enables departments with lower capacity to deliver practice-based DCE without relying on local champions.
- 2.4 Prioritise fewer, higher-leverage laboratories rather than many small additions. Where resources are constrained, focus on a small number of high-impact practical experiences (data acquisition → cleaning → modelling → interpretation → governance implications) supported by validated simulations when needed, but not as a substitute for hands-on work.

3. Build institutional capacity at scale: UK engineering higher education should build a community of practice to support teaching and learning in data engineering, including sector-wide micro-credentials on statistics for engineers, machine learning literacy, and the general growing use of AI in engineering education

- 3.1 A community of practice should be established to support the development and dissemination of professional development opportunities for academics and teaching staff – in statistics and data analytics for engineers, data engineering, ethics and governance around data and machine learning and AI literacy.
- 3.2 Create short, stackable CPD for academics and teaching staff. These nationally available micro-credentials should be aligned to the DCE outcomes learning outcomes.
- 3.3 Support new education-enabling roles and shared services. Encourage institutions to develop or share roles such as data stewards, learning technologists for data laboratories, and research software engineers (RSEs) who can sustain DCE laboratories and projects over time.
- 3.4 Target capacity investment to enable project-based delivery (compute, datasets, tooling). Where departments aspire to expand DCE provision, ensure minimum infrastructure (compute access, approved datasets, secure environments where needed, and licensing support). Without this, departments default to lecture-led provision.
- 3.5 Maintain a shared repository of exemplars, project briefs, assessment patterns, governance templates, and evaluation instruments, enabling rapid diffusion and reducing duplication of effort.
- 3.6 While not covered in this report, the Community of Practice could also provide a space for continued discussions around the growing use of generative AI and large language models (LLMs) more generally in engineering education (teaching and assessment).

4. Accreditation alignment: propose explicit data engineering language in Engineering Council accreditation (AHEP 5) and other skills standards, using national exemplars to guide programmes

- 4.1 Professional bodies and sector partners should explore explicit data engineering language in the next iteration of the Engineering Council accreditation standard for higher education, AHEP 5 for Bachelor's (level 6) and Integrated (MEng) and standalone (MSc) Master's (level 7). This reduces variability and signals that DCE is a standard feature of modern engineering education.
- 4.2 Require constructive alignment: outcomes → teaching → assessment evidence. Encourage programmes to evidence DCE coverage through mapped outcomes, module learning activities, and assessment blueprints, enabling consistent QA and comparability across the sector.
- 4.3 A suite of national exemplars (teaching and assessment materials) should be developed, created through the Community of Practice to support this embedding.
- 4.4 Where possible, the learning outcomes included in this report should also be aligned with the recently produced UK Standard Skills Classification.

5. Ensure inclusion and widen participation in DCE pathways

- 5.1 Adopt “on-ramp” supports to prevent DCE widening attainment gaps. Implement diagnostics and just-in-time support for maths, coding, and data handling; design inclusive tasks and communication that broaden participation, particularly in programmes where student preparedness varies.
- 5.2 Make DCE examples diverse, societally relevant, and sustainability-linked. Use context-rich engineering case studies (infrastructure resilience, decarbonisation, safety-critical systems) to motivate learning and support engagement across student cohorts.

6. Enable sector coordination and a targeted implementation pathway

- 6.1 Use a maturity pathway model: Emerging → Adopters → Innovator. Using this model to support a staged approach aligned to the institutional typologies identified in the analysis:
 - Type A Emerging:* enable minimum viable DCE (core outcomes + starter packs)
 - Type B Adopters:* deepen practice-based learning (laboratories/projects)
 - Type C Innovators:* scale exemplars and mentor others

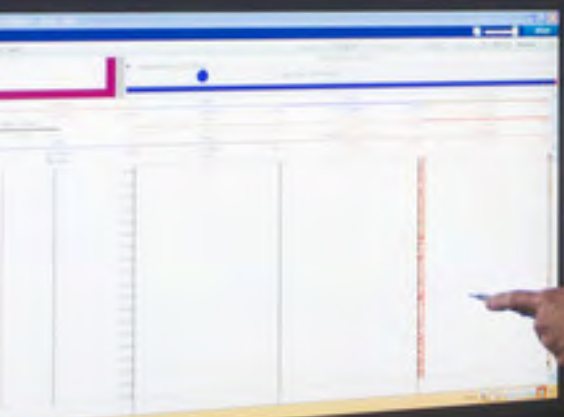
7. Update teaching, learning and assessment to reflect DCE practice and use of artificial intelligence tools

- 7.1 Redesign assessment to foreground judgement, explanation, and critique. Prioritise assessment of modelling choices, assumptions, uncertainty, ethical trade-offs, and interpretation rather than only procedural calculation.
- 7.2 Implement clear guidance on AI tools in teaching, learning and assessment, with consistent safeguards. Move from ad hoc departmental experimentation to sector-consistent practices: permitted uses, required disclosure, and assessment designs that test understanding and decision-making. This addresses the emerging reality that many departments are exploring or already allowing AI tools.



0.0	No Data			
No Data	No Data			
0	No Data	133		
2	0.00	1	0.0	0.00
0	0.00	11.0	-39.6	0.00

18696.0	0.0			
17320.0				
0	0	102		
-24	0	0	0.1	8.65
0	0	0.0	0.0	8.65



DD2

Annex A: Methodology

A.1 Introduction

The methodology was guided by the following considerations:

- **Research design:** We decide our study to be of mixed methods.
- **Data collection methods:** We conducted a desk research, surveys and semi-structured interviews.
- **Sampling strategy:** We collected data from 60 departments and conducted seven interviews.
- **Data analysis method:** We conducted quantitative and thematic analysis to establish a baseline map of DCE provision and explore qualitative insights.
- **Ethical considerations:** Consideration was given to GDPR and to participant consent and anonymity of interviewees and their institutions.

A.2 Preliminary

As a first step, a steering group comprised of experts and users of data engineering across different engineering disciplines and from academia and industry was formed to advise on the content and direction of this study. The group's work was key in all the stages of this initiative. The steering group hypothesised a set of learning outcomes based the learning pyramid (foundational, advanced, expert). The tree workshops took place and aimed to:

- identify learning outcomes across subject areas
- validate the identified Learning outcomes
- validate findings and co-construct practical recommendations in Phases III, IV.

Participants:

- academic leaders, professional bodies and industry partners.

Method:

- structured facilitation techniques (for example, nominal group technique)
- present preliminary findings from survey and interviews
- identify gaps, priority areas and recommendations for scaling DCE integration.

A.3 Methodological approach: Mixed methods

Embedding DCE in UK higher education involves both curricular structures (for example, how topics

are designed and delivered) and institutional strategies (for example, why and under what conditions they are adopted). To capture this dual dimension, we considered a mixed-methods design as most appropriate^{11,12}.

This approach allowed us to:

- map the *landscape of current DCE integration* (quantitative/documentary)
- understand motivations, barriers and innovations (qualitative)
- build an *evidence base* for policy and curriculum design.

A convergent parallel design was particularly suitable: we collected qualitative and quantitative data concurrently, analysed them separately and then integrated the findings for conclusions.

A.3.1 Documentary and curriculum analysis (Baseline mapping)

This stage is aimed at establishing what is currently taught and how, before any institutional engagement.

Data sources:

- publicly available degree specifications, module descriptors and course handbooks
- accreditation criteria from the Engineering Council and professional bodies.

A.3.2 Quantitative

We ran a nationwide survey to identify trends, common practices and institutional variations in embedding DCE. The rationale was that a survey would provide breadth and generalisability, identifying patterns across institutions and informing where in-depth qualitative exploration is most needed¹³.

Target respondents:

- programme leaders, teaching and learning leads in engineering faculties from the following departments: mechanical, civil, aeronautical, electrical/electronics, chemical engineering.

11 Creswell, J.W. & Plano Clark V.L. (2017). *Designing and Conducting Mixed Methods Research* (3rd edn). Thousand Oaks, CA: SAGE Publications.

12 Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th edn). Thousand Oaks, CA: Sage.

13 Bryman, A. (2016). *Social Research Methods* (5th edn). London: Oxford University Press.

Instrument design:

- include closed questions on teaching practices, assessment, skills emphasis and barriers
- add open-ended questions to capture context and innovation
- pilot the survey with a small sample to refine clarity and validity.

Dissemination:

- national circulation through networks such as the Engineering Professors' Council, social media (LinkedIn), HEIs networks, and relevant professional bodies
- use purposive sampling complemented by snowballing.

Because of the survey's structure (categorical and ordinal data from 60 departments), the conducive analyses were descriptive, crosstab supplemented by exploratory multivariate analysis to surface insights. This mixed-quantitative design aligns with best practices in education policy and curriculum research, supporting both sector-level benchmarking and evidence-informed recommendations for embedding DCE across UK higher education.

A. Descriptive

Purpose: to provide a national overview of how DCE topics (foundational statistics, advanced statistics, data management, data analytics, governance/ethics, and AI/ML) are taught across UK institutions. This establishes a baseline map of DCE provision across disciplines (mechanical, civil, electrical and so on) and enables visualisation (for example, bar or radar charts) of relative adoption rates¹⁴.

B. Cross-tabulations and chi-square tests

To identify associations between categorical variables, such as discipline type and DCE content delivery, or between university type and teaching coverage. This approach tests whether DCE inclusion varies systematically across disciplines or institutional characteristics¹⁵.

C. Multivariate analysis (exploratory factor and cluster analysis)

To identify latent dimensions and institutional typologies in DCE provision. This moves from descriptive to pattern recognition, revealing structural differences across the sector¹⁶.

A.3.3 Qualitative (Thematic analysis)

We aimed at gaining deeper insights into how or why DCE topics are integrated or not within curriculums. The semi-structured interviews provide rich, contextualised understanding that explains patterns found in the survey. They are particularly useful for exploring institutional strategies and pedagogical philosophies (Cohen et al., 2017).

Participants:

- academics designing and delivering relevant modules
- curriculum leads and senior teaching staff
- professional bodies.

Themes explored:

- curriculum design processes and rationales
- challenges (skills gaps, staff capacity, accreditation constraints)
- pedagogical innovations.

Mode:

- conduct interviews online
- use a semi-structured guide to balance consistency and flexibility.

A.3.4 Final workshop (Post-committee report)

This workshop supported knowledge co-production and policy relevance, ensuring findings are actionable and aligned with HEIs priorities^{17,18}.

Method:

- structured facilitation techniques (for example, nominal group technique)
- present preliminary findings from survey and interviews
- identify gaps, priority areas and recommendations for scaling DCE integration.

A.4 Limitations of the data

The survey provides useful early evidence but has some limitations, in scale. Only 60 of 400 contacted departments responded (approximately 15%), creating a substantial risk of nonresponse and self-selection bias: departments already engaged with DCE may be more likely to participate, inflating apparent levels of provision and capacity. In addition, the data are self-reported and collected at departmental level, so responses may reflect perceptions, vary by respondent role and cannot be independently verified against module specifications.

14 Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (8th edn). London: Routledge. <https://doi.org/10.4324/9781315456539>

15 Field, A.P. (2018) *Discovering Statistics Using IBM SPSS Statistics* (5th edn). Sage, Newbury Park.

16 Hair, J.F. et al. (2019). *Multivariate Data Analysis* (8th edn). Andover, Hampshire: Cengage Learning

17 Cornwall, A. & Jewkes, R. (1995) What Is Participatory Research? *Social Science and Medicine*, 41, 1667-1676. [http://dx.doi.org/10.1016/0277-9536\(95\)00127-S](http://dx.doi.org/10.1016/0277-9536(95)00127-S)

18 Bryson, J. M., Quick, K. S., Slotterback, C. S., & Crosby, B. C. (2013). Designing Public Participation Processes. *Public Administration Review*, 73, 23-34. <https://doi.org/10.1111/j.1540-6210.2012.02678.x>

Annex B: Results

B.1. Quantitative data analysis (survey)

B.1.1 Descriptive analysis

The descriptive analysis aimed at providing an overview of how the selected DCE topics (foundational statistics, advanced statistics, data management, data analytics, governance/ethics, and AI/ML) are taught across UK institutions. By doing so we established a baseline map of DCE provision across disciplines (mechanical, civil, electrical and so on). We computed the frequencies and percentages of responses for the proportion of students exposed (0–25%, 50%, 75%, all); the teaching modes (lectures, laboratories, projects); and the institutional capacity scores (1–5). Additionally, we calculated the central tendency (mean, median) and variability (SD, IQR) for capacity ratings. Results are presented in **Table B.1**.

A. Set up

- **Sample:** 59 departmental responses from UK engineering schools.
- **Disciplines:** Each department could tick multiple disciplines (General, Mechanical, Aerospace, Manufacturing, Civil, Electronic/Electrical, Chemical, Biomedical, Other).

Each discipline was treated as a Yes/No flag (e.g. “Mechanical: Yes” vs “Mechanical: No”) and tested its association with the outcome.

B. Level of DCE integration: For each department we counted the number of the seven DCE topics taught as core for all students:

“All students learn the majority of this content (i.e. it is core for all students)”

Across:

- Foundational statistics
- Advanced statistics
- Data engineering and management
- Data analytics and visualisation
- Data governance, ethics and communication
- ML and AI for data analytics
- Additional AI-related skills for engineering

Then we defined an overall DCE integration level:

- Low: 0–1 topics core
- Medium: 2–3 topics core
- High: 4–5 topics core

Distribution across the 59 departments:

- Low: 22 departments (37%)
- Medium: 32 departments (54%)
- High: 5 departments (9%)

For each topic, for example, ML, we consider “present” vs “none”.

- Present = anything other than “No students learn this content” in the ML question.
- None = “No students learn this content”.

Result: 51 departments (86%) have some ML content; 8 (14%) have none.

ML capacity (institutional). ML capacity ratings (1–5) recoded as:

- Low: 1–2
- Medium: 3
- High: 4–5

(7 departments did not give a capacity rating for ML.)

Teaching method for ML (from the ML section):

- Lectures (Yes/No)
- Laboratory classes (Yes/No)
- Project work (Yes/No)

Table B.1 shows a landscape where data-handling literacy (foundational statistics, analytics) is reasonably well embedded across UK engineering, but more advanced and infrastructural aspects of DCE (data engineering, sophisticated AI) are still concentrated in particular disciplines and often accessible only to subsets of students. There were a few outliers. For foundational statistics, the two ‘1’ ratings suggest a few smaller or nontraditional engineering departments with minimal in-house quantitative expertise. In the case of ML for example, the cluster of ‘1’ responses ($\approx 10\%$ of respondents) signals a national capacity gap, these departments likely have little or no staff experience or resource to teach ML-related topics. Findings of this analysis are summarised in **Table B.2**.

B.1.2 Cross-tabulations and chi-square tests

We conducted cross-tabulation tests to identify associations between categorical variables, such as discipline type and DCE content delivery. Mainly:

- chi-square tests of independence between:
 - discipline and level of DCE integration
 - teaching method and level of institutional capacity
- Cramér’s V test to measure association strength.

Table B.1 Summary of descriptive analysis of quantitative data

DCE Topic	Typical Student Exposure	% of Departments Teaching to All Students	% with No Exposure	Main Teaching Modes	Mean Capacity (1-5)	Capacity Trend	Key Insight
Foundational statistics	Very high (mean ≈ 89% of students)	74.6 %	0 %	Lectures (97%), Laboratories (54%), Projects (44%)	4.18 ± 1.15	Strong	Core element in almost all departments; high confidence and mature provision
Advanced statistics	Moderate (mean ≈ 39%)	13.6 %	15.3 %	Projects (37%), Lectures (31%), Laboratories (15%)	3.24 ± 1.12	Mixed	Taught mainly as optional or project content; patchy institutional expertise
Data engineering and management	Moderate – low (mean ≈ 36%)	8.5 %	20.3 %	Projects (34%), Lectures (29%), Laboratories (24%)	2.87 ± 1.18	Weak	Clear skills gap; limited curricular embedding and low staff capacity
Data analytics	High (mean ≈ 72%)	47.5 %	6.8 %	Projects (49%), Lectures (42%), Laboratories (41%)	3.80 ± 1.15	Good	Broadly taught and increasingly core; applied, project-based learning common
Data governance / ethics / professional skills	Moderate (mean ≈ 54%)	37.3 %	23.7 %	Lectures (27%), Projects (27%), Laboratories (7%)	3.22 ± 1.29	Variable	Often integrated into design or ethics modules; one-quarter report no coverage
ML and AI for data analytics	Moderate (mean ≈ 37%)	6.8 %	13.6 %	Projects (49%), Lectures (46%), Laboratories (31%)	3.35 ± 1.27	Moderate	Emerging presence, mainly in final year or specialist courses
Additional AI-related skills (e.g., NLP, vision, robotics AI)	Low (mean ≈ 28%)	5.1 %	22.0 %	Projects (41%), Lectures (25%), Laboratories (22%)	2.93 ± 1.20	Weak	Early-stage adoption; concentrated in computing-heavy disciplines

Table B.2 Summary of findings

Dimension	Finding
Coverage strengths	Foundational statistics and data analytics are near-universal, often compulsory for all students
Emerging areas	ML and governance/ethics are gaining ground but not yet standardised across programmes
Capacity gaps	Data management and advanced AI skills show low capacity and inconsistent provision
Pedagogical modes	Lectures dominate for core topics; project-based learning is common for advanced and applied topics
Disciplinary variation	Mechanical and electrical/electronic departments show highest adoption of analytics and ML; smaller disciplines remain uneven
AI tools readiness	68% of departments are exploring or already using AI tools in teaching; 47% allow use in assessment

Table B.3 Cross-disciplinary synthesis

Category	Strongest provision	Typical gaps	Distinctive features
High DCE Integration	Electronic/electrical, biomedical, mechanical	Data management infrastructure	Advanced ML & analytics incorporated in core modules
Moderate Integration	Civil, chemical, aerospace, general	Limited ML depth	Governance/ethics consistently included
Lower Integration	Manufacturing (small N)	Staff capacity & optional rather than core modules	Strong industrial/project focus

This was to test whether DCE inclusion varies systematically across disciplines. **Table B.3** presents the cross-disciplinary synthesis.

DCE inclusion does not systematically depend on discipline (chi-square not significant), but delivery depth and pedagogy do depend on institutional capacity, particularly for project-based ML and data management teaching. Strengthening staff skills and resourcing for applied data work will enable consistent DCE adoption across all branches of UK engineering.

B.1.3 Multivariate analysis (exploratory factor and cluster analysis)

We conducted a multivariate analysis to identify latent dimensions and institutional typologies in DCE provision. That means identifying the underlying dimensions that shape DCE provision and grouping UK engineering departments into distinct institutional typologies. First, we conduct the factor analysis.

Table B.4 presents the three-factor solution of the exploratory factor analysis, that is DCE provision in terms of three broad latent dimensions, namely:

- **Pedagogical integration**
How pervasively, across topics and modes, DCE appears in the curriculum.
- **DCE/AI capacity**
How confident departments feel about their ability to teach these topics, particularly advanced statistics and ML/AI.
- **Practice-based integration**
The extent to which DCE is delivered via projects and laboratories (especially analytics and governance), rather than only via lectures.

These factors give a compact, interpretable representation of the survey data, which we can then use to group institutions.

The cluster analysis surfaces three groups of institutions which contain a mix of integration levels in the simple 'core topics' sense, the structural differences residing more in how and with what capacity DCE is delivered.

Type A (Traditional / low-integration departments): These departments are traditional in their curriculum structures and appear to be at an early stage of DCE adoption. They have some capacity

Table B.4

Factor	Dominant loadings	Interpretation	Variance explained
F1 – Data pedagogical integration	High loadings for <i>lectures, laboratories, and projects</i> across foundational statistics, analytics and governance; moderate for advanced statistics	Reflects how broadly DCE concepts are embedded across teaching formats	34.2%
F2 – AI and advanced data capacity	Strong loadings on <i>capacity</i> for ML and additional AI Skills; moderate on advanced statistics capacity	Captures institutional depth in ML/AI capability and advanced analytics competence	22.8%
F3 – Data infrastructure and management readiness	Loadings on data management capacity + project-based delivery for data and analytics	Indicates data engineering resource and practice-based integration	14.9%

but relatively little DCE in the taught curriculum and project-based work.

Type B (Content-led DCE adopters): These departments have moved fastest to embed DCE into taught content, especially via lectures and structured modules. They are firmly ‘emerging adopters’: DCE is visible across the curriculum, but their capacity and practice-based integration lag behind the most advanced institutions.

Type C (High-capacity, practice-based DCE innovators): These are the leading-edge institutions: they combine strong capacity with a clear emphasis on authentic, project-based and lab-based DCE learning. Advanced statistics are not just talked about; they are applied in substantial project work.

These latest statistical tests, the factor and cluster analyses, bring the quantitative analysis beyond simple frequencies and percentages to reveal structural patterns in the provision of DCE.

- Three underlying dimensions organise DCE provision:
 - how widely DCE is built into the curriculum (Pedagogical Integration)
 - how confident departments are in delivering it (Capacity / AI capacity)
 - how far it is taught through authentic practice (Practice-based DCE).
- Three institutional archetypes emerge:
 - traditional departments – limited DCE presence, particularly ML, despite moderate capacity
 - content-led adopters – strong lecture-based DCE and ML; still developing capacity and practice-based work

- high-capacity innovators – strong capacity, and extensive project-/lab-based DCE, especially in ML and analytics.

B.1.4 Summary of quantitative analysis

This section synthesises all quantitative analyses in the survey of *Data-Centric Engineering (DCE) education in UK higher education*. **Table B.5** consolidates findings from the Descriptive analysis, Cross-tabulations and chi-square tests, and the Multivariate analysis (exploratory factor & cluster analysis), and provides a cross-validation commentary identifying where they align or diverge.

The three analytic layers reinforce one another, indicating a coherent national pattern:

- DCE foundations are in place, but advanced, applied data skills are confined to a smaller set of high-capacity innovators.
- Capacity – not discipline – is the core structural determinant.
- Departments evolve from lecture-based exposure (Pathway B) to project-based integration (Pathway C) as resources and confidence grow.
- Minor mismatches (for example, a few departments teaching ML despite low confidence, or disciplinary differences lacking statistical weight) reflect sampling limits, not contradictory trends.

B.2 Qualitative: Thematic analysis

B.2.1 Pedagogical integration: Embedding DCE

A dominant theme across all institutions is that interviewees reject a stand-alone, add-on model. There is a shared commitment to

Table B.5

Theme	Evidence base	Strategic implication
1. UK DCE education shows strong foundations but limited depth	High coverage in foundational statistics and analytics (descriptive + EFA)	Build beyond basics: national focus should shift from 'literacy' to 'applied data capability'
2. Institutional capacity, not discipline, drives DCE innovation	Confirmed across chi-square and multivariate results	Policy interventions should target capacity building rather than discipline-specific mandate.
3. Project-based learning is the hallmark of advanced DCE maturity	Shown by significant χ^2 for project work \times capacity and Factor 3 loadings	Incentivise practice-based curricula through funding and exemplars
4. Three institutional archetypes exist	Cluster analysis	Support pathways: (A) Introductory adoption, (B) Pedagogical enhancement, (C) Showcase exemplars
5. Cross-method coherence strengthens reliability	Triangulation across descriptive, inferential and multivariate results	Findings are robust, methodological convergence supports confidence in the baseline map

embedding DCE principles within existing engineering curriculums rather than creating stand-alone modules. DCE should be a thread that runs through laboratories, design projects and capstones with students encountering sensing, data acquisition, cleaning, visualisation, and interpretation in discipline-authentic contexts. HEI-B explicitly frames the aim as DCE “in students’ DNA”, embedded across projects and laboratories rather than discrete modules. Interviewees prioritise course redesign around authentic tasks – not new modules – so DCE becomes part of routine engineering judgement.

- HEI-C emphasised the integration of DCE through laboratory and design-based learning, for example embedding data handling and analysis in thermodynamics laboratories and Python-based projects (for example, flood monitoring).
- HEI-B described DCE as something that should be “in students’ DNA”, embedded through project-based learning and practical activities throughout the curriculum.
- HEI-A and HEI-D both reinforced this logic, seeing existing lab activities (thermodynamics, manufacturing, computational methods) as natural vehicles for embedding data collection, processing and analytics skills.

This reflects a pedagogical preference for integration that minimises curriculum overload, aligns with accreditation expectations and fosters more authentic learning experiences. This aligns with constructivist pedagogical approaches that

emphasise learning through application and problem-based learning frameworks¹⁹.

Proposal: Prioritise course redesign around existing tasks, not new modules, so DCE becomes part of routine engineering judgement.

B.2.2 Core v specialist competencies

A consensus model distinguishes foundational competencies for every degree student (data handling, elementary statistics, basic data engineering) from advanced competencies (for example, regression families, optimisation, ML/AI) for specialists or upper-year electives. This is articulated as a progression across undergraduate years and into Postgraduate. HEI-B explicitly outlined a structured learning progression: Years 1–2 for foundational skills (data handling, statistics, computational thinking) and Years 3–4 for discipline-specific applications. HEI-C similarly identified core competencies (for example, descriptive statistics, data acquisition) and advanced competencies (for example, ML, optimisation) and is aligning these with their new curriculum reform (*Design and Data-Centric Engineering* paper from 2027). HEI-A referred to Bloom’s modified taxonomy for engineering practice, differentiating between awareness, application and critical understanding while HEI-D

19 Hmelo-Silver, C.E., Eberbach, C. (2012). Learning Theories and Problem-Based Learning. In: S. Bridges, C. McGrath & T. Whitehill (eds) *Problem-Based Learning in Clinical Education. Innovation and Change in Professional Education*, Vol 8. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-2515-7_1

highlighted integration from early years through foundational computing and data science modules, then layering advanced methods in upper years.

Proposal: Adopt a two-tier outcomes framework, ensuring minimum DCE literacy for all while enabling specialisation through electives and projects.

B.2.3 Ethical, legal and societal dimensions of DCE

DCE is a natural vehicle to teach ethics, privacy, GDPR, bias, and responsible AI. Across all HEIs institutions, ethics and governance were seen as central, not peripheral to DCE education, in alignment with the growing “Responsible Engineering and Responsible AI agendas in UK higher education”. HEI-C argued for ethics to be embedded contextually “as professional reasoning”, rather than taught as an “add-on”, particularly around AI, privacy and responsible innovation. HEI-D treats ethical and professional considerations, including data privacy and accountability, as embedded learning outcomes from first year onwards. They embed ethical accountability from Year 1 and cautions against “black-box” use of ML. HEI-A raised issues related to AI in assessment and academic integrity, stressing the need for institutional guidance. HEI-B also flagged inclusion and accessibility as ethical considerations in framing DCE content.

Proposal: Define assessable learning outcomes on ethical data practice and uncertainty quantification across the spiral curriculum.

B.2.4 Institutional capacity, staff development and cultural change

Interviewees consistently noted that staff confidence and institutional readiness are critical to successful DCE implementation. All sites report variable staff confidence (especially in probability/statistics and ML) and highlight the need for CPD and shared resources. HEI-C highlighted uneven levels of statistical literacy among faculty and suggested reframing content around “uncertainty and risk” to build consensus. HEI-A pointed to capacity gaps in less data-intensive disciplines and emphasised the need for CPD resources and sector-wide support. HEI-B described DCE integration as a “cultural shift,” requiring faculty buy-in and sustained professional development. HEI-D noted the challenge of staying up to date with rapidly evolving software tools while focusing on fundamental skills. HEI-A and HEI-B stress uneven readiness across departments and the need for open CPD toolkits and practice communities.

Proposal: Pairing curricular change with systematic CPD, exemplars and coaching; community of practice on assessment redesign in an AI-pervasive environment.

B.2.5 Accreditation and policy levers

All institutions linked DCE adoption to the evolution of accreditation frameworks, particularly the AHEP.

HEI-C and HEI-A explicitly discussed the constraints and opportunities presented by AHEP 4, with optimism that AHEP 5 will allow greater emphasis on DCE. HEI-A proposes explicit data/AI references in outcomes (M1/M2 and C1/C2) and placement at FHEQ L6 to ensure universal proficiency.

HEI-B saw alignment with accreditation as essential for scaling adoption nationally. They positioned national guidance and a community of practice to influence AHEP 5 and normalise embedded DCE. HEI-D expressed willingness to engage with shaping AHEP 5 through early adoption and resource development. HEI-A noted that accreditation already privileges application (ethics, sustainability, simulation) over encyclopaedic science coverage, undermining ‘we can’t fit it in’ arguments.

These considerations highlight the role of regulatory and policy frameworks as enablers of pedagogical change, a finding consistent with previous national curriculum reform efforts in engineering education.

Proposal: Aligning programme learning outcomes to an AHEP-consistent DCE spine, using accreditation language to legitimate rebalancing.

B.2.6 Inclusivity and accessible entry routes

HEI-B highlighted inclusive framing to avoid discouraging students with weaker coding backgrounds; language matters for participation of under-represented groups. HEI-F interview added that uneven prior maths/practicals demand bridging toolkits and careful scaffolding.

Proposal: Provide on-ramp resources (diagnostic support, just-in-time maths/data ‘toolkits’) and inclusive task design from Year 1.

B.2.7 Resources, laboratories and scale: embedding under real constraints

- The HEI-F interview surfaced hard constraints, technician capacity, lab throughput (for example, 14 repeats), equipment at scale, post-COVID drift to “dataset-only” laboratories and warned that authentic hands-on is expensive yet critical to connect maths with engineering.

Proposal: Prioritise high-leverage laboratories (fewer, richer), instrument existing experiments for data work and support with virtualisation only where it complements but not replaces hands-on.

B.2.8 Software-agnostic fundamentals and assessment redesign

- HEI-C urges software-agnostic teaching (principles over brand-specific tools) and elevates Bayesian reasoning/uncertainty. HEI-D echoes focus on transferable modelling and critical evaluation rather than tool-chasing. HEI-A and HEI-F emphasise assessment challenges in the LLM era and the need to evidence authentic engineering reasoning.

Proposal: Redesign assessments to foreground judgement, explanation and validation; permit tool use but assess understanding and decision-making.

B.2.9 Industry alignment and employability

HEI-E underscores employer demand for DCE skills and shows project-led mastery (for example, data cleaning for industrial rigs), while acknowledging uneven depth across student choices; staff capability often tracks research strengths. HEI-B links DCE to employability narratives and recommends national showcases and early-adopter pathways.

Proposal: Expand industry-anchored projects and visiting-professor partnerships to keep tasks authentic and skill signals legible to employers.

B.2.10 Points of divergence (productive tensions)

Across interviewees, the emerging commonality from a strategic standpoint is that DCE is framed as a transformative lens. That means, a way engineers think and decide under uncertainty, rather than an extra topic to add in the curriculum. HEI-C's forthcoming *Design and Data-Centric Engineering* core paper epitomises this reframing. The HEI-D redesign similarly integrates data/science/ethics from Year 1. DCE is treated as a capability architecture that re-organises existing teaching, enabling simplification (not accretion) of curricula.

However, diverging views have also surfaced:

- Where to locate statistics/ML depth? Some favour widespread light-touch exposure; others prefer concentrated advanced options. (HEI-C and HEI-E v mixed views at HEI-A).
- Space v substitution. HEI-E argues for taking content out to make room for DCE; others prefer embedding within existing laboratories with minimal timetable change.
- Facilities realism. HEI-F cautions that scaling authentic laboratories is hard without technician and kit investment; this may temper ambitious embedding plans.

These are summarised in **Table B.6**.

Summary of the matrix

- Themes 1 (**Pedagogical integration**), 2 (**Scaffolding**), and 10 (**Strategic framing**) are strongly shared across all six institutions, indicating broad convergence on treating DCE as an embedded, transformative capability rather than a stand-alone subject.
- Themes 4 (**Staff development**) and 5 (**Accreditation**) are identified as critical enablers, without investment in faculty capability and AHEP evolution, implementation may stall.
- Theme 6 (**Inclusivity**) shows differentiated emphasis, with HEI-B taking a leadership stance, signalling a potential gap in sector-wide practice.
- Theme 7 (**Resources**) is most strongly articulated by HEI-F, suggesting that operational and infrastructural constraints may not be adequately addressed elsewhere.
- Theme 8 (**Assessment redesign**) is widely acknowledged, reflecting the growing awareness of how LLMs and AI tools affect engineering education.
- Theme 9 (**Industry alignment**) is present across interviews, but operationalised differently, ranging from explicit projects (HEI-E) to strategic visions (HEI-C, HEI-B).

Table B.6 Matrix showing views of different HEIs for each theme

Themes	HEI-A	HEI-B	HEI-C	HEI-D	HEI-E	HEI-F
1. Pedagogical integration (embedding, not adding)	✓ Integration in laboratories	✓ Strong – “in students’ DNA,” project-based	✓ Strong – Python, laboratories, flood monitoring	✓ Laboratories, modelling focus	✓ Strong – project-anchored	✓ Highlights lab constraints but supports embedding
2. Curricular scaffolding (core v advanced tiers)	✓ Bloom’s ladder mapping	✓ Progressive model	✓ Distinct core v advanced	✓ Scaffolding noted	✓ “Baseline for all” + advanced pathways”	✓ Emphasises need for structured progression
3. Ethics and governance (responsible DCE)	✓ Ethical implications in assessment	✓ Ethical framing linked to inclusion	✓ Integrated ethics	✓ Embedded from Year 1	✓ Implied in industry links	– Limited reference
4. Staff development & capacity	✓ Noted as major barrier	✓ CPD required for culture shift	✓ Uneven stats literacy	✓ Challenge of tool evolution	✓ Linked to research strengths	✓ Critical constraint
5. Accreditation & policy levers (AHEP)	✓ Explicit proposals for AHEP 5	✓ Framing DCE via AHEP 5	✓ Linked to curriculum reform	✓ Supportive	✓ Leverages flexibility of AHEP	✓ Sees as needed enabler
6. Inclusivity and accessibility	–	✓ Strong – inclusive framing	– Minimal emphasis	–	–	✓ Bridging tools for uneven skills
7. Resources and lab scale	–	–	–	–	–	✓ Emphasises technical, logistical constraints
8. Software-agnostic fundamentals & assessment redesign	✓ Assessment in AI era	✓ Emphasised tool-agnosticism	✓ Bayesian reasoning focus	✓ Modelling principles	✓ Tool use contextualised	✓ Highlights AI-assessment issues
9. Industry alignment & employability	✓ Accreditation link	✓ Strategic framing	✓ Vision for graduate capability	✓ Curriculum reform alignment	✓ Industry projects central	✓ Resource implications
10. Strategic framing of DCE (transformative lens)	✓ Accreditation leadership	✓ Sector leadership	✓ Core paper	✓ Curriculum redesign	✓ Project-led model	✓ Realism on scale and constraints

✓ = theme present

Strong = explicitly and extensively emphasised

– = limited or no explicit reference





Royal Academy of Engineering

The **Royal Academy of Engineering** creates and leads a community of outstanding experts and innovators to engineer better lives. As a charity and a Fellowship, we deliver public benefit from excellence in engineering and technology and convene leading businesspeople, entrepreneurs, innovators and academics from every part of the profession. As a National Academy, we provide leadership for engineering and technology, and independent, expert advice to policymakers in the UK and beyond.

We have three goals:



Sustainable and Innovative Economy, where sustainability drivers, innovative industries and resilient infrastructures are aligned to drive growth and productivity that will support better lives for all.



Technology Improving Lives, where technology in all its forms is used to meet the most important human needs, avoid harm, support fairer societies and break down barriers to opportunity.



Engineering Community Fit for the Future, where our community reflects society in its diversity, commits to creating inclusive cultures to help drive engineering excellence, and has the skills to meet future needs safely, securely and ethically, and to keep pace with innovation.

Our work is enabled by funding from the Department for Science, Innovation and Technology, corporate and university partners, charitable trusts and foundations, and individual donors.

Royal Academy of Engineering
Prince Philip House
3 Carlton House Terrace
London SW1Y 5DC

Tel: +44 (0)20 7766 0600
www.raeng.org.uk

Registered charity number 293074