Module Descriptor

Title	Data Engineering		
Session	2025/26	**Status**	Published
Code	COMP10085	**SCQF Level**	10
Credit Points	20	**ECTS (European Credit Transfer Scheme)**	10
School	**Computing, Engineering and Physical Sciences**		
Module Co-ordinator	TBC		

Summary of Module

Data engineering has emerged as a critical discipline in modern software system design. With the exponential growth of the amount of data collected and processed, this discipline addresses data challenges ranging from scalability and consistency to efficiency and maintainability of data infrastructure and pipelines. In today's data-rich environment, professionals across various domains, including software engineering, data science, and analytics need to have a comprehensive understanding of data engineering practices. As a role, data engineers are responsible for the design and management of the infrastructure and pipelines that collect, transform, and store data so that it can then be used in downstream tasks by data analysts, data scientists, business intelligence professionals, and other stakeholders.

This module aims to equip students with the skills to plan and construct data systems that effectively meet organizational and customer needs. By exploring the data engineering lifecycle, students will gain insights into selecting and integrating diverse technologies to cater to downstream data consumers.

The syllabus will include the following topics:

- Data Engineering Overview
- Definition, Roles, Activities, and Responsibilities
- Introduction to the Data Engineering Lifecycle
- Phases: Generation, Storage, Ingestion, Transformation, Data Serving
- Data Architecture
- Principles
- Concepts: Domains and Services, Distributed Systems, Coupling
- Types: Data Warehouse, Data Lake, Data Mesh
- Data Engineering Life Cycle
- Data Generation Sources: APIs, Files, Databases
- Data Formats: JSON, XML, and Language-Specific Formats
- Data Storage Systems: File, Block, Object, Cache, Streaming Storage

- Data Storage Abstractions: Data Warehouse, Lake, Lakehouse, OLTP, OLAP
- Data Ingestion: Key Considerations (Frequency, Sync/Async, Throughput, Reliability, Security), Batch and Message/Stream Ingestion, Key Methods (APIs, Databases, SSH, SFTP, Queue/Event-Streaming)
- Building Data Pipelines
- ETL and ELT Overview
- Extracting Data from Various Sources (NoSQL, SQL DB, APIs)
- Loading Data into Data Warehouses
- Transforming and Validating Data
- Best Practices
- Queries, Modeling, and Transformation
- Queries: Definition, Lifecycle, Queries on Streaming Data
- Data Model: Definition, Types, Modeling Streaming Data
- Transformations: Batch Transformation, Materialized Views, Query Virtualization, Streaming Transforms, Processing
- Data Serving
- Considerations: Trust, Use Cases, Users, Products
- Analytics: Business, Operational, Embedded
- Ways to Serve Data: File Exchange, Database, Streaming
- Security and Privacy
- People, Processes, and Technology to Enforce Security and Privacy

Throughout the practical sessions of this module, students will develop and deploy components of data pipelines across the data engineering lifecycle into suitable cloud, on-premises, or local systems. They will also explore the usage of various mainstream industry tools that enable them to implement the core principles of data engineering.

This module will work to develop a number of the key 'I am UWS' Graduate Attributes to make those who complete this module:

- Universal: critical thinker; analytical; and research-minded
- Work Ready: motivated; digitally literate; effective communicator; and ambitious
- Successful: innovative; resilient; and transformational

Module Delivery Method	On-Campus[1] ☐	Hybrid[2] ☒	Online[3] ☐	Work -Based Learning[4] ☐
Campuses for Module Delivery	☐ Ayr ☐ Dumfries	☒ Lanarkshire ☐ London ☐ Paisley	☒ Online / Distance Learning ☐ Other (specify) Online Delivery / Distance Learning applies to delivery in the BSc (Hons) Data, AI and Software Engineering programme only	

Terms for Module Delivery	Term 1	☒	Term 2	☐	Term 3	☐
Long-thin Delivery over more than one Term	Term 1 – Term 2	☐	Term 2 – Term 3	☐	Term 3 – Term 1	☐

Learning Outcomes	
L1	Demonstrate a detailed understanding of the fundamental principles and concepts of data engineering, including the roles, activities, and responsibilities of a data engineer
L2	Explain the phases of the data engineering lifecycle, such as data generation, storage, ingestion, transformation, and data serving, and critically appraise the key considerations and technologies associated with each phase
L3	Develop and deploy data pipelines using suitable tools and principles of data engineering given a practical scenario
L4	Critically evaluate and compare different data storage systems, ingestion methods, and data serving techniques, considering factors such as reliability, security, and scalability and analyse the effectiveness of various data engineering approaches
L5	N/A

Employability Skills and Personal Development Planning (PDP) Skills	
SCQF Headings	**During completion of this module, there will be an opportunity to achieve core skills in:**
Knowledge and Understanding (K and U)	**SCQF 10**

[1] Where contact hours are synchronous/ live and take place fully on campus. Campus-based learning is focused on providing an interactive learning experience supported by a range of digitally-enabled asynchronous learning opportunities including learning materials, resources, and opportunities provided via the virtual learning environment. On-campus contact hours will be clearly articulated to students.

[2] The module includes a combination of synchronous/ live on-campus and online learning events. These will be supported by a range of digitally-enabled asynchronous learning opportunities including learning materials, resources, and opportunities provided via the virtual learning environment. On-campus and online contact hours will be clearly articulated to students.

[3] Where all learning is solely delivered by web-based or internet-based technologies and the participants can engage in all learning activities through these means. All required contact hours will be clearly articulated to students.

[4] Learning activities where the main location for the learning experience is in the workplace. All required contact hours, whether online or on campus, will be clearly articulated to students

	A critical understanding of the principal theories, concepts and principles in the domain of data engineering. Detailed knowledge and understanding of how to create and assess a data pipeline.
Practice: Applied Knowledge and Understanding	**SCQF 10** Apply knowledge, skills and understanding in executing the development and implementation of data pipelines along the lifecycle of data engineering.
Generic Cognitive skills	**SCQF 10** Identify problems, analyse results, and interpret common error messages to solve problems in a logical manner Consolidate knowledge, concepts, skills, practices and methodology in the production of data engineering solutions.
Communication, ICT and Numeracy Skills	**SCQF 10** This subject area is entirely computer based so ICT skills feature heavily in the practice of the subject area. Communicating effectively and appropriately in commercial style, numerate, written reports produced using standard office ICT equipment and software.
Autonomy, Accountability and Working with Others	**SCQF 10** Exercise autonomy and initiative to implement independently parts of data engineering solutions at a professional level

Prerequisites	**Module Code**	**Module Title**
	Other	
Co-requisites	**Module Code**	**Module Title**

Learning and Teaching	
In line with current learning and teaching principles, a 20-credit module includes 200 learning hours, normally including a minimum of 36 contact hours and maximum of 48 contact hours.	
Learning Activities During completion of this module, the learning activities undertaken to achieve the module learning outcomes are stated below:	**Student Learning Hours** (Note: Learning hours include both contact hours and hours spent on other learning activities)
Lecture / Core Content Delivery	18
Tutorial / Synchronous Support Activity	6
Laboratory / Practical Demonstration / Workshop	24
Independent Study	152
Please select	
Please select	

TOTAL	

Indicative Resources
The following materials form essential underpinning for the module content and ultimately for the learning outcomes: Bryan, K. and Ransome, T. (2023) Cracking the Data Engineering Interview. Packt Publishing Ltd. Crickard, P. (2020) Data Engineering with Python. Packt Publishing Ltd. Densmore, J. (2021) Data Pipelines Pocket Reference. O'Reilly Media. Kleppmann, M. (2017) Designing Data-Intensive Applications. 'O'Reilly Media, Inc.' Reis, J. and Housley, M. (2022) Fundamentals of Data Engineering. 'O'Reilly Media, Inc.'
(N.B. Although reading lists should include current publications, students are advised (particularly for material marked with an asterisk*) to wait until the start of session for confirmation of the most up-to-date material)

Attendance and Engagement Requirements
In line with the Student Attendance and Engagement Procedure, Students are academically engaged if they are regularly attending and participating in timetabled on-campus and online teaching sessions, asynchronous online learning activities, course-related learning resources, and complete assessments and submit these on time. **For the purposes of this module, academic engagement equates to the following:** The School of Computing, Engineering and Physical Sciences considers attendance and engagement to mean a commitment to attending, and engaging in, timetabled sessions. You will scan your attendance via the scanners each time you are on-campus and you will login to the VLE several times per week. Where you are unable to attend a timetabled learning session due to illness or other circumstance, you should notify the Programme Leader that you cannot attend. Across the School an 80% attendance threshold is set. If you fall below this, you will be referred to the Student Success Team to see how we can best support your studies.

Equality and Diversity
The University's Equality, Diversity and Human Rights Procedure can be accessed at the following link: UWS Equality, Diversity and Human Rights Code. Aligned with the University's commitment to equality and diversity, this module supports equality of opportunity for students from all backgrounds and learning needs. Using the VLE, material will be presented electronically in formats that allow flexible access and manipulation of content. This module complies with University regulations and guidance on inclusive learning and teaching practice. This module has lab-based teaching and as such you are advised to speak to the Module Co-ordinator to ensure that specialist assistive equipment, support provision and adjustment to assessment practice can be put in place, in accordance with the University's policies and regulations.
(N.B. Every effort will be made by the University to accommodate any equality and diversity issues brought to the attention of the School)

Supplemental Information

Divisional Programme Board	**Computing**

Overall Assessment Results	☐ **Pass / Fail** ☒ **Graded**
Module Eligible for Compensation	☐ **Yes** ☒ **No** **If this module is eligible for compensation, there may be cases where compensation is not permitted due to programme accreditation requirements. Please check the associated programme specification for details.**
School Assessment Board	Business & Applied Comupting
Moderator	Graeme McRobbie
External Examiner	A Jindal
Accreditation Details	
Module Appears in CPD catalogue	☐ **Yes** ☒ **No**
Changes / Version Number	1.01

Assessment (also refer to Assessment Outcomes Grids below)
Assessment 1
Assessment 1: Written report analysing a case study/problem, requiring students to describe and propose data engineering solutions aligned with the lifecycle. (40%)
Assessment 2
Assessment 2: A coursework assessment that requires a group of students to act as a development team and develop, document, and evaluate a data engineering pipeline on a cloud platform, demonstrating their practical application of data engineering concepts and proficiency in using taught technologies. (60%)
Assessment 3
(N.B. (i) Assessment Outcomes Grids for the module (one for each component) can be found below which clearly demonstrate how the learning outcomes of the module will be assessed. (ii) An indicative schedule listing approximate times within the academic calendar when assessment is likely to feature will be provided within the Student Module Handbook.)

Component 1							
Assessment Type	**LO1**	**LO2**	**LO3**	**LO4**	**LO5**	**Weighting of Assessment Element (%)**	**Timetabled Contact Hours**
Review/ Article/ Critique/ Paper	☒	☒	☐	☐	☐	40	

Component 2							
Assessment Type	**LO1**	**LO2**	**LO3**	**LO4**	**LO5**	**Weighting of Assessment Element (%)**	**Timetabled Contact Hours**
Portfolio of practical work	☐	☐	☒	☒	☐	60	

Component 3							
Assessment Type	**LO1**	**LO2**	**LO3**	**LO4**	**LO5**	**Weighting of Assessment Element (%)**	**Timetabled Contact Hours**
	☐	☐	☐	☐	☐		
Combined total for all components						100%	hours

Change Control

What	**When**	**Who**
Attendance and Engagement Procedure and Equality and Diversity	17/1/25	F.Valentine