



## **Module Specification**

### **Big Data [TSI]**

Version: 2023-24, v2.0, 06 Dec 2023

#### **Contents**

<b>Module Specification .....</b>	<b>1</b>
<b>Part 1: Information .....</b>	<b>2</b>
<b>Part 2: Description .....</b>	<b>2</b>
<b>Part 3: Teaching and learning methods .....</b>	<b>4</b>
<b>Part 4: Assessment.....</b>	<b>5</b>
<b>Part 5: Contributes towards .....</b>	<b>6</b>

## Part 1: Information

**Module title:** Big Data [TSI]

**Module code:** UFCEE1-12-M

**Level:** Level 7

**For implementation from:** 2023-24

**UWE credit rating:** 12

**ECTS credit rating:** 6

**College:** College of Arts, Technology and Environment

**School:** CATE School of Computing and Creative Technologies

**Partner institutions:** Transport and Telecommunication Institute

**Field:** Computer Science and Creative Technologies

**Module type:** Module

**Pre-requisites:** None

**Excluded combinations:** None

**Co-requisites:** None

**Continuing professional development:** No

**Professional, statutory or regulatory body requirements:** None

## Part 2: Description

**Overview:** This module will provide students an understanding of concepts and techniques which underpin the capture, management and interrogation of large data sets.

**Features:** Not applicable

**Educational aims:** The aim of this module is to provide students with an understanding of technologies, challenges and trends within the subject area of Big Data and Cloud Computing

**Outline syllabus:** Data Storage and Retrieval:

Importance of data for business.

Understand the difference between data, information and knowledge.

Traditional ways to store and retrieve data.

Big Data challenges and opportunities.

Introduction to Big Data:

Defining Big Data: Sources of Big Data. The four dimensions of Big Data - volume, velocity, variety, veracity. Introducing storage and MapReduce.

Business application of Big Data: Big Data applications/examples in business.

Delivering business benefit from Big Data. Establishing the business importance of Big Data.

Addressing the challenge of extracting useful data/knowledge.

Integrating Big Data with traditional data.

SQL Databases vs. NoSQL Databases:

Understand the growing amounts of data.

The relational database management systems (RDBMS).

Capabilities of traditional RDBMSs.

Overview of Structured Query Languages (e.g. SQL).

Introduction to NoSQL databases.

Understanding the difference between a relational DBMS and a NoSQL database.

Identifying the need to employ a NoSQL DB.

Storing Big Data:

Analysing data characteristics: Selecting data sources for analysis.

Introduction of selected Big Data stores from the following list: Hadoop, Cassandra, Amazon S3, BigTable, etc.

Achieving Data Quality:

Introduction to data quality.

Why is data quality a business problem?

Problems when data is not “fit for purpose”.

Preparing data.

Ways to improve data quality.

Understand ETL - Extract, Transform, Load procedures to improve Data Quality.

Knowledge-based Information Retrieval:

Introduction to knowledge-based information retrieval.

Use for ontologies for knowledge modelling.

Learn how to build an ontology to link knowledge with data.

Using ontologies for information retrieval – case study.

Machine learning for knowledge acquisition: Introduction to machine learning and pattern recognition. Capabilities of different modelling, analysis and algorithmic techniques.

Big Data and Cloud Computing (technology, challenges and trends):

Cost of storing Big Data.

Is cloud computing a solution?

Issues: privacy and trust.

Future of Big Data and cloud computing.

Future research trends in Big Data.

### **Part 3: Teaching and learning methods**

**Teaching and learning methods:** The module is delivered through weekly lectures and weekly tutorial sessions. Each lecture will direct the course and introduce the new ideas and skills required. Then small group tutorial sessions will enable each student to carry out the study and research exercises described in the associated work-sheet under the guidance of a Tutor.

**Module Learning outcomes:** On successful completion of this module students will achieve the following learning outcomes.

**MO1** Describe similarities and differences between big data lifecycle frameworks and propose correct methods for each stage for successful integration.

**MO2** Apply suitable analysis methods from data identification, acquisition, pre/processing of different data types, modelling and evaluation to performing error analysis and optimisation.

**MO3** Present a comprehensive plan for handling each stage of the big data lifecycle that justifies choice of methods.

**MO4** Design and develop data products, derived from successfully integrated methods from machine learning, data visualization and UI/UX.

**Hours to be allocated:** 120

**Contact hours:**

Independent study/self-guided study = 112 hours

Face-to-face learning = 48 hours

Total = 160

**Reading list:** The reading list for this module can be accessed at [readinglists.uwe.ac.uk](https://rl.talis.com/3/uwe/lists/55CF09B9-8FC4-6BF1-4426-90839BFCA98B.html?lang=en-gb&login=1) via the following link <https://rl.talis.com/3/uwe/lists/55CF09B9-8FC4-6BF1-4426-90839BFCA98B.html?lang=en-gb&login=1>

## Part 4: Assessment

**Assessment strategy:** The assessment is a project including a viva.

Project (individual completion).

This task involves students investigating a business related cloud adoption problem based on given requirements, proposing a solution and preparing implementation specifications. The actual assignment topics are carefully chosen to demonstrate some basic principles, which are especially significant to the course.

Oral presentation/Viva (individual completion).

This task consists of questions related to course work produced by the student

which should test the students understanding of the concepts presented in the course work as well as their understanding and ability to apply those concepts and ideas to real-life scenarios (case studies).

There will be opportunities for formative assessment in the form of regular in-class presentations of research/implementation completed as part of tutorial work completed and subsequent group discussions.

The resit is a rework/update of the project.

**Assessment tasks:**

**Project (First Sit)**

Description: Written Report (3,000 words) plus Viva (15 mins)

Weighting: 100 %

Final assessment: No

Group work: No

Learning outcomes tested: MO1, MO2, MO3, MO4

**Project (Resit)**

Description: Written Report (3,000 words) plus Viva (15 mins)

Weighting: 100 %

Final assessment: No

Group work: No

Learning outcomes tested: MO1, MO2, MO3, MO4

**Part 5: Contributes towards**

This module contributes towards the following programmes of study:

Computer Science (Data Analytics and Artificial Intelligence) {Double Degree} [TSI]

MSc 2023-24