

MODULE SPECIFICATION

Part 1: Basic Data					
Module Title	Big Data				
Module Code	UFCF8H-15-M	Level	M		
Owning Faculty	FET	Field	Computer Science and Creative Technologies		
Contributes towards	MSc Information Technology MSc Information Management MSc Financial Technology				
UWE Credit Rating	15	ECTS Credit Rating	7.5	Module Type	Standard
Pre-requisites	None		Co- requisites	None	
Excluded Combinations	None		Module Entry requirements	None	
Valid From	September 2019				

CAP Approval Date	February 2015 v1, November 2015 v1.1, June 2016 v1.2, UVP 29 May 2019 v2
--------------------------	--------------------------------------------------------------------------

Part 2: Learning and Teaching	
Learning Outcomes	<p>On successful completion of this module students will be able to:</p> <ol style="list-style-type: none"> 1. Understand the importance of data for business applications and the difference between data, information and knowledge in terms of their uses [A] 2. Understand the challenges in storage and retrieval of small and large amounts of data, and the difference between SQL and NoSQL databases [A] 3. Apply problem solving skills necessary for identifying the organizational needs to employ a SQL or NoSQL DB [B] 4. Understand the four dimensions of Big Data i.e. volume, velocity, variety, veracity, which are important challenges the delivery of business benefits from Big Data [A] 5. Be able to apply problem-solving skills to address the challenge of extracting useful data and application of data quality checks [A, B] 6. Master various ways to improve data quality by understanding why data quality is a business problem [A, B] 7. Apply knowledge modelling skills to generate ontologies to define domain knowledge and relationships between entities, and use them for information retrieval purposes [B] 8. Demonstrate knowledge of Big Data management using Cloud computing and associated privacy and trust issues [A, B]
Syllabus Outline	<p>Data Storage and Retrieval</p> <ul style="list-style-type: none"> • 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 <p>Introduction to Big Data</p> <ul style="list-style-type: none"> • Defining Big Data

STUDENT AND ACADEMIC SERVICES

	<ul style="list-style-type: none"> ○ Sources of Big Data ○ The four dimensions of Big Data: Volume, velocity, variety, veracity ○ Introducing storage and MapReduce ● Business application of Big Data <ul style="list-style-type: none"> ○ 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 <p>SQL Databases vs. NoSQL Databases</p> <ul style="list-style-type: none"> ● 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 <p>Storing Big Data</p> <ul style="list-style-type: none"> ● Analysing data characteristics <ul style="list-style-type: none"> ○ Selecting data sources for analysis ● Introduction of selected Big Data stores from the following list: <ul style="list-style-type: none"> ○ Hadoop, Cassandra, Amazon S3, BigTable, etc. <p>Achieving Data Quality</p> <ul style="list-style-type: none"> ● 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 <p>Knowledge-based Information Retrieval</p> <ul style="list-style-type: none"> ● 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 <ul style="list-style-type: none"> ○ Introduction to machine learning and pattern recognition ○ Capabilities of different modelling, analysis and algorithmic techniques <p>Big Data and Cloud Computing (technology, challenges and trends)</p> <ul style="list-style-type: none"> ● 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 												
Contact Hours	<p>This module will involve 2 hours contact time per week for one semester equally divided between lecture and tutorial sessions.</p> <table border="1" data-bbox="587 1854 1270 2047"> <thead> <tr> <th>Activity</th> <th>hrs</th> </tr> </thead> <tbody> <tr> <td>Contact time</td> <td>24</td> </tr> <tr> <td>Assimilation and development of knowledge</td> <td>86</td> </tr> <tr> <td>Exam preparation</td> <td>10</td> </tr> <tr> <td>Coursework preparation</td> <td>30</td> </tr> <tr> <td>Total study time</td> <td>150</td> </tr> </tbody> </table>	Activity	hrs	Contact time	24	Assimilation and development of knowledge	86	Exam preparation	10	Coursework preparation	30	Total study time	150
Activity	hrs												
Contact time	24												
Assimilation and development of knowledge	86												
Exam preparation	10												
Coursework preparation	30												
Total study time	150												

STUDENT AND ACADEMIC SERVICES

<p>Teaching and Learning Methods</p>	<p>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.</p> <p>The teaching material is available from Blackboard. A course text is also recommended.</p> <p>Scheduled learning includes lectures and tutorials.</p> <p>Independent learning includes time engaged with essential reading and assignment preparation and completion.</p>
<p>Reading Strategy</p>	<p>Core readings This is an emerging field within information science discipline; consequently there are no specific textbooks that can adequately cover all the module contents. Students will therefore be directed to read specific chapters and/or sections from the listed reference books and articles throughout the course of the module. These will be made available via library/blackboard.</p> <p>Further readings Students are also expected to identify all other reading relevant to their independent research topic for themselves. They will be encouraged to read widely using the library search, a variety of bibliographic and full text databases, and Internet resources. Many resources can be accessed remotely.</p> <p>Access and skills The development of literature searching skills is supported by a Library seminar provided within the first semester. These level three skills will build upon skills gained by the student whilst studying at levels one and two. Additional support is available through the Library Services web pages, including interactive tutorials on finding books and journals, evaluating information and referencing. Sign-up workshops are also offered by the Library.</p> <p>Indicative reading list The following list is offered to provide validation panels/accrediting bodies with an indication of the type and level of information students may be expected to consult. As such, its currency may wane during the life span of the module specification. However, as indicated above, <i>current</i> advice on readings will be available via the module guide.</p>
<p>Indicative Reading List</p>	<p>Books: Roebuck, K. (2012) <i>Storing and Managing Big Data – NoSQL, Hadoop and more</i>. Emereo Publishing [ISBN: 1743045743].</p> <p>Davis, K. (2012) <i>Ethics of Big Data</i>. O'Reilly Media [ISBN:978-1-4493-1178-0]</p> <p>Glenn J. Myatt, Wayne P. Johnson, (2014), <i>Making Sense of Data, 2nd Edition A Practical Guide to Exploratory Data Analysis and Data Mining</i>, Publisher Wiley [ISBN: 978-1-118-40741-7]</p> <p>Articles: Labrinidis, A. and Jagadish, H. V. (2012), <i>Challenges and opportunities with big data</i>. <i>Proc. VLDB Endow.</i> 5, 12,, pp. 2032-2033.</p> <p>Madden, S. (2012) <i>From Databases to Big Data, Internet Computing, IEEE</i> , vol.16, no.3, pp.4,6.</p>

STUDENT AND ACADEMIC SERVICES

Part 3: Assessment	
Assessment Strategy	<p>The assessment is split 25/75 between an oral assessment presentation/viva exam (Component A) and coursework (Component B). The Component B, coursework normally involves solving 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.</p> <p>The coursework is required to be carried out by individual students and the assessment should be made on written work provided by each individual.</p> <p>The component A grade will be obtained from a single oral assessment presentation/viva exam to be taken at the completion of teaching. This component will consist of questions related to course work produced by the student which should test the students understanding of the fundamental 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).</p> <p>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.</p>

Identify final assessment component and element	Component A	
% weighting between components A and B (Standard modules only)	A: 25%	B: 75%
First Sit		
Component A (controlled conditions) Description of each element	Element weighting (as % of component)	
1. Oral Assessment Presentation/Viva Exam (10-15 Minutes)	100%	
Component B Description of each element	Element weighting (as % of component)	
2. Individual Course Work Report (2000 Words)	100%	

Resit (further attendance at taught classes is not required)		
Component A (controlled conditions) Description of each element	Element weighting (as % of component)	
1. Oral Assessment Presentation/Viva Exam (10-15 Minutes)	100%	
Component B Description of each element	Element weighting (as % of component)	
2. Individual Course Work Report (2000 Words)	100%	