UNIVERSITY OF LONDON

BSc Computer Science

Module Specification

Key Information			
Module title	Databases and Advanced Data Techniques		
Level	6	Credit value	15
Member Institution	Goldsmiths	Notional study hours and duration of course	150
Module lead author/ Subject matter expert	Dan McQuillan		
Module co-author			

Rationale for the module

Databases and data services form the basis of many of the computer systems we use today. The collection, storage, filtering and processing of data is, therefore, a critical application area for computer science. This module builds on data concepts and skills from previous modules, providing advanced level knowledge about working with data in computers.

Aims of the module

This module aims to show you how to work with data in your computer programs. You will learn how to use SQL and NoSQL databases to store tabular data and documents. You will learn about the ethics of gathering and processing data and why it is important to consider issues around data security. You will learn about open data resources, and how you can access them from your computer programs. You will learn about audio and video data, and the challenges of working with this kind of data.

Topics covered in this module:

The topics listed here are an approximation of what will be covered. The topics presented may be slightly revised to ensure currency and relevance. Students will be advised of any changes in advance of their study.

1. Open data sources: different data formats
2. Gathering and cleaning data
3. The ethics of working with data
4. SQL: advanced data representation
5. SQL: advanced queries and linking
6. NoSQL and document stores
7. NoSQL and data pipelines
8. Speeding up queries with indexes
9. Audio and video data
10. Data security

Approximately 10-12 hours of study will be required per topic. The remaining study time is intended for coursework and examination preparation.

Learning outcomes for the module

Students who successfully complete this module will be able to:

1. Devise and explain an appropriate representation of data in an SQL RDBMS
2. Write advanced SQL queries using indexes and joins to read and write data from an SQL database
3. Write data to a NoSQL database and selectively read data from a NoSQL database using filters
4. Read data from various file formats and databases and prepare it for use in a computer program using data processing techniques
5. Discuss critical issues relating to the ethics of working with data and data security and apply this knowledge to particular situations
6. Understand the challenges of working with audio and video data and explain how systems for working with this data overcome these challenges

Assessment strategy, assessment methods

Summative and Formative Assessments
The module will contain a range of summative and formative assessments. Summative assessments are assessments which contribute directly towards your final grade. Formative assessments do not count directly towards your final grade. Instead, they provide you with opportunities for low stakes practice, and will often provide some sort of feedback about your progress. For example, a practice quiz might provide you with feedback about why a particular answer was wrong.

Assessment Activities
The table below lists the assessment activity types you might encounter taking the module. It also states if that type of assessment can be automatically graded. For example, multiple choice quizzes can be automatically graded, and so can some programming assignments. It also states if that type of assessment will be found in the summative coursework and the summative examination. More details about the summative assessments are provided below.

Assessment activity type	Can it be automatically graded with feedback in some cases?	Coursework	Examination
Quiz	X	X	X
Writing task		X	X
Programming task	X	X	X
Peer review task		X	

Pass Mark

In order to pass this module, you must achieve at least 35% in each element of summative assessment and an overall weighted average of 40%, subject to the application of rules for compensation. Please refer to the programme regulations for more information.

Summative Assessment Elements

As this is a module that has a significant amount of theory it is assessed as a theory-based module. This means that the summative assessment is composed of two elements, whose weightings are listed in the table below.

Summative Assessment Component	Percentage of final credit	Deadline
Coursework	50%	Mid session
Examination	50%	End of session

The coursework comprises a variety of practical exercises and quizzes which in total will take up to 25 hours of study time to complete. The examination will be two hours long, and consist of written answer and multiple choice questions.

Learning resources

The module will draw on a number of different, largely web-based, public resources as well as the resources produced as bespoke material for this module. The standard text book(s) for the module will be:

C J Date. Database Design and relational Theory. O'Reilly 2012

Connolly and Begg, *Database Systems*. Pearson 2016