



Department of Computer Science and
Information Systems

MSc in Data Science
Programme Arrangements
2017–2018

Version of September 12, 2017

Contents

1	General Information	4
1.1	Contacts	4
1.2	Web presence	4
2	Student Support	5
3	Important Dates	6
3.1	Term Dates	6
4	Syllabus	7
4.1	Introductory Talks	7
4.2	List of Modules	7
4.2.1	Compulsory Modules	7
4.2.2	Optional Modules	8
4.3	Timetables	9
5	Module Descriptions	11
5.1	Fundamentals of Computing (FoC)	11
5.2	Principles of Programming I (PoP I)	13
5.3	Big Data Analytics Using R (BDA)	15
5.4	Computer Systems (CS)	17
5.5	Programming with Data (PwD)	19
5.6	Data Science Techniques and Applications (DSTA)	21
5.7	Cloud Computing (CC)	23
5.8	Information and Network Security (INSEC)	25
5.9	Data Warehousing and Data Mining (DWDM)	27
5.10	Information Retrieval and Organisation (IRO)	29
5.11	Information Systems (IS)	31
5.12	Machine Learning (ML) — formerly Intelligent Technologies	33
5.13	Principles of Programming II (PoP II)	35
5.14	Semantic Technologies (ST)	37
5.15	Advances in Data Management (ADM)	39
6	MSc Project	41
7	Administration and Assessment	43
7.1	Requirements for the Award of the MSc	43
7.2	Announcement of Results	44
7.3	Choosing the Optional Modules	44
7.4	Mitigating Circumstances and Deferral	44

7.5	Retake and Reassessment	45
7.6	Re-enrolment	46
7.7	Examinations	47
7.8	Coursework	47
7.9	Projects	48
7.10	Assessment Offences and Plagiarism	48
8	Student Services	49
8.1	Counselling Service	49
8.2	Mental Health Service	49
8.3	Disability and Dyslexia Service	49
9	The Business Engagement Team	51
9.1	Mentoring Pathways	51
9.2	Enterprise Pathways	51
9.3	Keeping in Touch	51
10	Career Development	52

1 General Information

1.1 Contacts

Programme Director:	Dr Alessandro Provetti,	ale@dcs.bbk.ac.uk
Programme Administrator:	Zahra Syed,	pgadmin@dcs.bbk.ac.uk
Admissions Tutors:	Dr Carsten Fuhs,	carsten@dcs.bbk.ac.uk
	Dr David Weston,	dweston@dcs.bbk.ac.uk
Projects Tutor:	Dr Oded Lachish,	oded@dcs.bbk.ac.uk
Disability Officer:	Dr Oded Lachish,	oded@dcs.bbk.ac.uk

1.2 Web presence

Detailed and updated information about the programme is available from the

- internet page:

<http://www.dcs.bbk.ac.uk/study/postgraduate/msc-data-science/>

- intranet page (for enrolled students):

www.dcs.bbk.ac.uk/dcswiki/index.php/MSc_DS_Intranet

The Moodle Virtual Learning Environment

moodle.bbk.ac.uk

is used to provide detailed information and post announcements about each module on which you are enrolled.

It is your responsibility to familiarise yourself with the contents of both of this booklet as well as the programme's web site, and to consult the web site on a regular basis, since additional information will be posted there during the year. You should also read your College email on a regular basis.

2 Student Support

Every student is allocated a personal tutor in the first weeks of the programme. The personal tutor is someone students can contact to discuss any problems of a non-academic nature. These may relate to special needs or personal problems that may affect the student's academic performance. The Department also has a disability officer whom students can contact.

Academic problems should first be addressed to the lecturer concerned. If the problem is not resolved or it does not relate to a specific module, then the Programme Director should be contacted.

For more general information about Birkbeck, student services and regulations have a look at

`http://www.bbk.ac.uk/student-services`.

It is expected that students familiarise themselves with these pages so that they are aware of the services and regulations.

The School of Business, Economics and Informatics has Learning Co-ordinators. Their role is to support students in their studies. They can offer help and support on a variety of topics ranging from writing skills to basic maths. See

`www.bbk.ac.uk/business/current-students/learning-co-ordinators`

for details.

3 Important Dates

3.1 Term Dates

Lectures will commence in the week starting on Monday 2nd October 2017. The teaching (i.e., not including exams and project) covers two terms of eleven weeks each (autumn and spring term). The summer term is given over to revision (including revision lectures), exams, and the project.

- Autumn term: Monday 2nd October 2017 – Friday 15th December 2017.
- Spring term: Monday 8th January 2018 – Friday 23rd March 2018.
- Summer term: Monday 23rd April 2018 – Friday 6th July 2018.

Please refer to

<http://www.bbk.ac.uk/about-us/term-dates>

for the College holiday closing times.

Students should attend lectures during term time as shown in the timetables in Section 4.3. If students are unable to attend lectures, they should arrange with lecturers or fellow-students to obtain copies of any material distributed in class.

Any student who decides to withdraw from the programme should inform the Programme Administrator, in writing or by email. Students who simply stop turning up for lectures without formally withdrawing from the programme will still be held liable for fees. It is especially important for international students that they inform the department about any absence.

4 Syllabus

4.1 Introductory Talks

The programme will kick off with introductory talks to new students:

- Part-time students: 18:00, Thursday September 28th 2017, room MAL 405;

These will include a short hands-on introduction to the departmental computer system. There will also be short presentations by representatives of the library and the disability office.

Background material

On the course web site enrolled students will find three notes, in PDF format, to be used for pre-course preparation. The three notes will refresh students on Python programming, Mathematics and Probability & Statistics, respectively.

4.2 List of Modules

We give a general overview of the content of the programme here; detailed description of the individual modules is in the next section.

4.2.1 Compulsory Modules

- Fundamentals of Computing — FoC (15 credits)
- Principles of Programming I — PoP I (15 credits)
- Big Data Analytics using R — BDA (15 credits)
- Computer Systems — CS (15 credits)
- Programming with Data — PwD (15 credits)
- Data Science Techniques and Applications — DSTA (15 credits)
- MSc Data Science Project (60 credits)

4.2.2 Optional Modules

Students enrolled in 2017-18 have to choose two optional modules from the list below. Students enrolled in 2016-17 have to choose only one module for the same list.

- Cloud Computing — CC (15 credits)
- Information and Network Security — INSEC (15 credits)
- Data Warehousing and Data Mining — DWDM (15 credits)
- Information Retrieval and Organisation — IRO (15 credits)
- Information Systems — IS (15 credits)
- Machine Learning — ML (15 credits)
- Principles of Programming II — PoP II(15 credits)
- Semantic Technologies — ST (15 credits)
- Advances in Data Management — ADM (15 credits)

Additionally, a 15-credit, level 7 module offered on the Advanced Computing Technologies MSc programme. For a detailed description of these modules and the timetable, see the following web page:

<http://www.dcs.bbk.ac.uk/courses/mscact/>

Please note that the modules above run on a different timetable, so you can only take these modules if you are able to fit them into your schedule. If you are interested in taking a module from the Advanced Computing Technologies MSc programme as an option, please discuss this with the MSc DS Programme Director.

Optional module availability is subject to timetabling constraints and student demand. In the event that an optional module is over-subscribed, available places will be allocated on a first-come, first-served basis determined by the date you return your module choice form to the Programme Administrator.

4.3 Timetables

The teaching venues will be announced online:

<http://www.dcs.bbk.ac.uk/courses/>

For an overview of the teaching venue locations, please refer to

<http://www.bbk.ac.uk/mybirkbeck/guides/help/class-information/>

Below is the timetable for the modules. Note that occasionally there might be changes (e.g. swapping of lectures between modules, or additional tutoring sessions). Please contact regularly the web pages of the modules for up-to-date information.

Part-time Students, Year 1

In the first year, Data Science students take only compulsory modules as follows.

Term 1 (Autumn)	18:00–19:20	19:40–21:00
Mon.	PoP I	PoP I
Tue.	BDA	BDA
Wed.	FoC	
Thu.		
Fri.		

Term 2 (Spring)	18:00–19:20	19:40–21:00
Mon.		
Tue.	PwD	PwD
Wed.		FoC
Thu.	CS	CS
Fri.		

Part-time Students, Year 2

In their second year, PT students enrolled in 2016-17 take the compulsory modules *Big Data Analytics using R (BDA)* and *Data Science Techniques and Applications*

(*DSTA*), and also one optional module.

Term 1 (Autumn)	18:00–19:20	19:40–21:00
Mon.		
Tue.	BDA	BDA
Wed.		
Thu.		
Fri.		

Term 2 (Spring)	18:00–19:20	19:40–21:00
Mon.		
Tue.		
Wed.	DSTA	DSTA
Thu.		
Fri.		

5 Module Descriptions

5.1 Fundamentals of Computing (FoC)

Teaching Staff

Michael Zakharyashev (module coordinator), Trevor Fenner

Online material

<http://www.dcs.bbk.ac.uk/~michael/foc/foc.html>

<http://www.dcs.bbk.ac.uk/~trevor/FoC/focTF.html>

Aims

Discrete mathematics, mathematical logic, and the related fundamental areas of data structures and algorithms lie at the heart of any modern study of Computer Science. Any understanding of how computers operate and how to use them effectively and efficiently, in terms of either their hardware or software, inevitably involves numerous mathematical concepts.

Syllabus

- Digital logic. Arithmetic for computers.
- Elements of set theory.
- Finite state machines (automata). Nondeterministic automata.
- Regular languages.
- Context-free languages and pushdown automata.
- Turing machines. Universal Turing machines. Undecidable problems.
- Data structures: representations and operations.
- Lists, stacks, queues and dequeues.
- Trees, forests, binary trees.
- Tree traversal and other operations; binary search trees.
- Organisation of disk storage; methods of file organisation; B-trees.
- Design and analysis of algorithms. Sorting and searching.

Prerequisites

Students taking this module must be also be currently taking (or have previously taken) a suitable programming module (Principles of Programming, or Introduction to Software Development). With the permission of the Programme Director, other students may take this module if they have equivalent appropriate programming experience.

Assessment

By 3-hour written examination and coursework exercises, weighting 80% and 20% respectively.

Reading

- D. Patterson and J. Hennessy, Computer Organization and Design: The Hardware/Software Interface. Morgan Kaufmann; 3 edition, 2007.
- E. Kinber and C. Smith, Theory of Computing. A gentle introduction. Prentice Hall, 2001.

5.2 Principles of Programming I (PoP I)

Teaching Staff

Keith Mannock (module co-ordinator), Vladislav Ryzhikov

Online material

On Moodle.

Aims, Outline and Learning Outcomes

To provide the student with a comprehensive grounding in programming.

This module introduces programming concepts and techniques, as well as elementary software development principles. Both for absolute beginners and for those with prior programming experience, the module introduces the fundamentals of object-oriented programming, including: classes and objects, variables and assignment, primitive and object types, methods, control structures, collections, iteration and recursion. The module also introduces basic software development issues such as class design, testing, debugging and documentation.

On successful completion of this module a student will be expected to be able to:

- Demonstrate knowledge of fundamental imperative programming concepts such as variables and assignment, conditional statements, loops and methods.
- Code an application in a suitable object-oriented programming language, from a detailed software specification.
- Evidence knowledge of basic concepts and principles of object-orientation such as objects and classes, encapsulation, object state, coupling, cohesion and modularity.
- Show awareness of basic principles of software design and development including appropriate naming of variables and classes, code layout, testing and debugging, and documentation.
- To show awareness of working in pairs and groups, and noting when, and where, they are most relevant.

Syllabus

- Core imperative programming ideas: sequence, selection, iteration, assignment, and variables.

- Version Control
- Data types
- Arrays
- Functions
- Automated testing and test driven development (TDD)
- Object-oriented programming: Encapsulation, Inheritance, Polymorphism, Message passing
- Dynamic data structures: Linked lists, Queues, and Stacks
- Recursion and recursive data structures
- Exception handling
- Program design in the small and the large

Prerequisites

A pass in the relevant admissions test.

Assessment

By examination (80%) and programming coursework (20%). The examination is split into two parts; 1) 40% via an online programming exam, and 2) 40% via a written exam.

Reading

A set of course notes will be provided but the following text covers similar topics:

- Python for Everyone, Interactive Edition, 2nd Edition, by Cay S. Horstmann, Rance D. Necaise. Wiley. 2016. ISBN: 978-1-119-18665-6

5.3 Big Data Analytics Using R (BDA)

Teaching Staff

Tingting Han

Online material

<http://moodle.bbk.ac.uk>

Aims

This module covers the principle concepts and techniques of data analytics and how to apply them to large-scale data sets. Students develop the core skills and expertise needed by data scientists, including the use of techniques such as linear regression, classification and clustering. The module will show you how to use the popular and powerful data analysis language and environment R to solve practical problems based on use cases extracted from real domains.

Syllabus

- Introduction to big data analytics: big data overview, data pre-processing, concepts of supervised and unsupervised learning.
- Basic statistics: mean, median, standard deviation, variance, correlation, covariance.
- Linear regression: simple linear regression, introduction to multiple linear regression.
- Classification: logistic regression, decision trees, SVM.
- Ensemble methods: bagging, random forests, boosting.
- Clustering: K-means, K-medoids, Hierarchical clustering, X-means.
- Evaluation and validation: cross-validation, assessing the statistical significance of data mining results.
- Selection of advanced topics such as: scalable machine learning, big data related techniques, mining stream data, social networks.
- Tools: R.

Prerequisites

Experience with a modern programming language.

Assessment

By 3-hour written examination and practical coursework, weighting 80% and 20% respectively.

Reading

- Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani. An Introduction to Statistical Learning: With Applications in R. Springer, 2017. Available online at <http://www-bcf.usc.edu/~gareth/ISL/>.

5.4 Computer Systems (CS)

Teaching Staff

Szabolcs Mikulás

Online material

<http://www.dcs.bbk.ac.uk/~szabolcs/compsys.html>

Aims

To learn the basics of computer architecture and organisation, and the role and mechanism of operating systems.

Syllabus

- Introduction: Computer architecture (CA) and Operating system (OS) overview
- Processors
- Processes and threads
- Concurrency
- Memory management
- I/O and file systems
- Protection and security
- Distributed and parallel processing

Prerequisites

MSc IT students who wish to enrol to this module are expected to pass the test for the short course.

Assessment

By 2-hour written examination and coursework, weighting 90% and 10%, respectively.

Reading

Textbook:

- W. Stallings, Operating Systems, Internals and Design Principles, Prentice Hall, 5th (or later) edition, 2005.

Recommended reading:

- W. Stallings, Computer Organization and Architecture: Designing for Performance, Prentice Hall, 7th (or later) edition, 2006.
- A.S. Tanenbaum, Modern Operating Systems, Prentice Hall. 2nd (or later) edition, 2001.

5.5 Programming with Data (PwD)

Teaching Staff

Alessandro Proveti

Online material

<http://www.dcs.bbk.ac.uk/~ale/dsta/>

<https://bitbucket.org/ale66/>

Aims

This module builds on the core modules “Big Data Analytics using R” and “Principles of Programming I” to further the students ability to program in Python, by introducing the specific Python modules that are designed for dataset processing. Moreover, it introduces the Relational Data Model and the SQL language for managing and querying databases. Finally, students will study and discuss the Ethics of Computing and Data Mining.

Syllabus

- introduction to the Python modules that are most commonly in use for Data Science, eg. numpy, pyplot, pandas and scikit;
- introduction to the Relational Data Model, and SQL as a language for data manipulation/querying, and
- the Ethical issues with Computing and Data collection and storage.

Prerequisites

The ability to program in Python (such as obtained by having taken the module Principles of Programming I, or as approved by the module leader).

Assessment

By 2-hour written examination and practical coursework, weighting 80% and 20%, respectively.

Reading

- Cay S. Horstmann, Rance D. Nicaise, Python for Everyone, Interactive Edition (2nd ed.). Wiley, 2016.

- J. Grus, Data Science from Scratch – First principles with Python. O'Reilly, 2015.
- Materials on SQL, TBC.

5.6 Data Science Techniques and Applications (DSTA)

Teaching Staff

Alessandro Provetti

Online material

<http://www.dcs.bbk.ac.uk/~ale/dsta/>

<https://bitbucket.org/ale66/>

Aims

This module provides students with more advanced study of data analytics, focussing on a range of applied data analysis techniques to convert information into knowledge.

Syllabus

- introduction to the module; definitions of Data Science;
- statistics and probability refresher;
- the relational data model and how to query SQL databases;
- experiences with Data Science discovery in Python;
- Web data extraction;
- from data to graphs, and their relevant properties;
- centrality measures;
- communities;
- correlation (if time allows).

Prerequisites

The ability to program in Python, SQL and a basic knowledge of statistics (such as obtained by having taken the module Programming with Data or as approved by the module leader).

Assessment

By 2-hour written examination and practical coursework, weighting 80% and 20%, respectively.

Reading

- J. Grus, Data Science from Scratch – First principles with Python. O'Reilly, 2015.
- G. Caldarelli and A. Chessa, Data Science and Complexity Networks. Oxford University Press, 2016.
- S. Marsland, Machine Learning – An Algorithmic Perspective (2ed). CRC Press, 2015.

Elective Modules Descriptions

5.7 Cloud Computing (CC)

Teaching Staff

Dell Zhang

Online material

<http://www.dcs.bbk.ac.uk/~dell/teaching/cc/>

Aims

This module aims to introduce back-end cloud computing techniques for processing “big data” (terabytes/petabytes) and developing scalable systems (with up to millions of users). We focus mostly on MapReduce, which is presently the most accessible and practical means of computing for “Web-scale” problems, but will discuss other techniques as well.

Students in this module will learn to understand the emerging area of cloud computing and how it relates to traditional models of computing, and gain competence in MapReduce as a programming model for distributed processing of big data.

Syllabus

- Introduction to Cloud Computing
- Cloud Computing Technologies and Types
- Big Data
- MapReduce and Hadoop
- Running Hadoop in the Cloud (Practical Lab Class)
- Developing MapReduce Programs
- Data Management in the Cloud
- Information Retrieval in the Cloud
- Link Analysis in the Cloud
- Beyond MapReduce
- Selected Case Studies
- Advanced Topics in Cloud Computing

Prerequisites

Good knowledge of Java programming is necessary. Students who did not have much experience in this area before joining their respective MSc programmes should have already taken the Introduction to Software Development module or Principles of Programming I and II.

Assessment

By 2-hour written examination and practical coursework, weighting 80% and 20%, respectively.

Reading

- J. Rosenberg and A. Mateos, *The Cloud at Your Service*, Manning, 2010.
- J. Lin and C. Dyer, *Data-Intensive Text Processing with MapReduce*, Morgan and Claypool, 2010.
- Extensive use is made of other relevant book chapters and research papers that are distributed or provided online.

5.8 Information and Network Security (INSEC)

Teaching Staff

David Weston

Online material

<http://www.dcs.bbk.ac.uk/~dweston/>

Aims

Information security is about protecting information (and information systems) against unauthorised access and tampering. Avoiding security breaches has a high priority for organisations storing and handling confidential data.

The main aim of this module is to provide broad coverage of the field of information security. This course covers the technical as well as the management side of security in information systems. Despite being an essential part of security, technical methods such as cryptography are not enough to guarantee a high level of security. They have to be embedded into a wider context in order to make them more effective. Users of technology have to understand the underlying principles and follow certain policies to avoid security breaches. This module introduces the fundamental approaches to security engineering and includes a detailed look at some important applications.

Syllabus

- Overview of Information Security
- Access Control Matrix Model
- Security Policies
- Social Engineering
- Basic Cryptography
- Identity Management
- Access Control Mechanisms
- Confinement
- Assurance and Trust
- Network Intruders and Intrusion Detection
- Firewalls and Malicious Software
- Cryptographic Protocol Concepts
- Authentication
- Key Exchange
- Economics of Information Security

Assessment

Two-hour written examination (80%) and practical coursework (20%).

Reading

- Keith M. Martin, *Everyday Cryptography: Fundamental Principles and Applications*, 2012, ISBN 978-0-19-969559-1
- William Stallings, *Cryptography and Network Security* 5th edition, Pearson, 2010, ISBN 0136097049
- Matt Bishop, *Computer Security: Art and Science*, Addison-Wesley, 2002, ISBN 0201440997
- Ross Anderson, *Security Engineering* 2nd edition, John Wiley & Sons, 2008, ISBN 0470068523
- Bruce Schneier, *Applied Cryptography*, John Wiley & Sons, 1996, ISBN 0-471-11709-9

5.9 Data Warehousing and Data Mining (DWDM)

Teaching Staff

Nigel Martin

Module URL

<http://www.dcs.bbk.ac.uk/~nigel/teaching/dwdm/>

Aims and Outline

To study advanced aspects of data warehousing and data mining, encompassing the principles, research results and commercial application of the technologies.

This module covers the organisation, analysis and mining of large data sets to support business intelligence applications. Students study the principles and commercial application of the technologies, as well as research results and emerging architectures underpinning the analysis and mining of “big data”.

Syllabus

- Data warehousing requirements.
- Data warehouse conceptual design.
- Data warehouse architectures.
- Data warehouse logical design: star schemas, snowflake schemas, fact tables, dimensions, measures.
- OLAP architectures, OLAP operations. SQL extensions for OLAP.
- Data warehouse physical design: partitioning, parallelism, compression, indexes, materialized views, column stores.
- Data warehouse construction: data extraction, transformation, loading and refreshing. Warehouse metadata. Continuous ETL.
- Data warehouse architecture trends. MapReduce and warehouse architectures: Pig, Hive, Spark.
- Data mining concepts, tasks and algorithms.
- Data mining technologies and implementations. Techniques for mining large data sets, stream mining, architecture trends, standards, products.
- Research trends in data warehousing and data mining.

Prerequisites

A first module in Database Systems (e.g. as taught in a typical UK undergraduate degree in computer science) or the Birkbeck module Data and Knowledge Management.

Assessment

By 2-hour written examination and practical coursework, weighting 90% and 10% respectively.

Reading

- R. Ramakrishnan, J. Gehrke, Database Management Systems (3rd ed.), McGraw Hill, 2003, ISBN 0-07-246563-8.
- M. Golfarelli, S. Rizzi, Data Warehouse Design: Modern Principles and Methodologies, McGraw Hill, 2009, ISBN 978-0-07-161039-1.
- J. Celko, Joe Celko's Analytics and OLAP in SQL, Morgan Kaufmann, 2006, ISBN 978-0-12-369512-3.
- J. Han, M. Kamber, J Pei, Data Mining Concepts and Techniques (3rd ed.), Morgan Kaufmann, 2011, ISBN 978-0-12-381479-1.
- Research papers will be distributed to students; students will also be directed to Web resources on the subject.

5.10 Information Retrieval and Organisation (IRO)

Teaching Staff

Dell Zhang

Online material

<http://www.dcs.bbk.ac.uk/~dell/teaching/ir/>

Aims and Outline

The aim of this module is to introduce modern Information Retrieval (IR) concepts and techniques, from basic text indexing to advanced text mining and Web IR. Both theoretical and practical aspects of IR systems will be presented and the most recent issues in the field of IR will be discussed. This will give students an insight into how modern search engines work and are developed.

Due to the explosive growth of digital information in recent years, modern Information Retrieval (IR) systems such as search engines have become more and more important in almost everyone's work and life (e.g. see the phenomenal rise of Google). IR research and development are one of the hottest research areas in academia as well as industry. This module will convey the basic principles of modern IR systems to students.

Syllabus

- Boolean Retrieval
- The Term Vocabulary and Postings Lists
- Regular Expressions and Text Normalization
- Dictionaries and Tolerant Retrieval
- Edit Distance
- Index Compression
- Scoring, Term Weighting and the Vector Space Model
- Evaluation in Information Retrieval
- Probabilistic Information Retrieval
- Language Models for Information Retrieval
- Language Modeling with N-Grams
- Spelling Correction and the Noisy Channel
- Text Classification
- Naive Bayes

- Sentiment Classification
- Vector Space Classification
- Flat Clustering
- Hierarchical Clustering
- Vector Semantics
- Semantics with Dense Vectors
- Matrix Decompositions and Latent Semantic Indexing

Assessment

By 2-hour written examination and practical coursework, weighting 80% and 20%, respectively.

Reading

- Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, Introduction to Information Retrieval, Cambridge University Press, 2008, ISBN 0521865719.
<http://www-csli.stanford.edu/~hinrich/information-retrieval-book.html>
- Dan Jurafsky and James H. Martin, Speech and Language Processing, 3rd ed draft. <https://web.stanford.edu/~jurafsky/slp3/>

5.11 Information Systems (IS)

Teaching Staff

Brian Gannon

Online materials

At the college Moodle server (i.e. at <http://moodle.bbk.ac.uk/>).

Aims and Outline

The primary aim of the module is to help students understand how information and communications technologies are deployed and to make informed professional decisions about IS development in fast changing socio-technical environments. This includes understanding how to use information processing constructs including files and data schemata, programs and other coded units, and the contexts to which they will be fitted. A subsidiary aim is to introduce students to some of the practical aspects associated with a career as an IS professional, and to describe key social and organisational aspects of enterprise computing.

The module describes approaches, processes, methodologies and techniques commonly used for large-scale information systems development. It covers the systems development life cycle (SDLC), including project initiation, analysis, design and implementation, addressing key aspects and techniques at each stage, such as the use of class and object diagrams. Several project methodologies are described, including Agile (Scrum) methodology. The module also incorporates insights into professional and legal issues surrounding Information Systems development.

Syllabus

- Theories of Information Systems
- The Information Systems Development Life Cycle
- Project Initiation - Identification and Selection
- Requirements Analysis & High Level Design
- Detailed design, including Architecture, DB design and UI design
- IS implementation
- Agile methodologies
- IS in everyday life
- IS contracts and legal issues
- Data Protection, Freedom of Information and Intellectual Property Rights
- Computer Misuse and Information Surveillance

Assessment

By 2-hour written examination and in-class tests, weighting 80% and 20%, resp.

Reading

- Tegarden, Dennis and Wixom, Systems Analysis and Design with UML, International Student Version, latest edition.
- Other supplementary readings will be advised.

5.12 Machine Learning (ML) — formerly Intelligent Technologies

Teaching Staff

George Magoulas

Online material

<http://moodle.bbk.ac.uk/>

Aims

The module covers computational algorithms for intelligent information management, decision making and complex problem solving. It provides an introduction to machine learning methods such as neural networks, fuzzy logic, fuzzy clustering, natural computing, and covers basic concepts of feature selection and generalisation.

Syllabus

- Learning from data
- Feature selection and generalisation
- Supervised learning, unsupervised learning and clustering
- Fuzzy logic and fuzzy clustering
- Deep networks: architectures and learning algorithms
- Natural computing: genetic algorithms, evolutionary algorithms, swarm intelligence
- Advanced learning and evolution schemes

Prerequisites

No specific module is pre- or co-requisite but knowledge of mathematical concepts (algebraic concepts, vector, matrix, function and graph, gradient, trigonometry concepts, statistical concepts and the notion of probability), and data structures and algorithms, as taught in a typical undergraduate degree in computer science or engineering, is essential.

Assessment

By 2-hour written examination and practical coursework, weighting 80% and 20%, respectively.

Reading

- S. Marsland, Machine Learning: an Algorithmic Perspective, CRC Press.
- M. Negnevitsky, Artificial Intelligence: a Guide to Intelligent Systems, Addison Wesley.
- E. Alpaydin, Introduction to Machine Learning, MIT Press.
- Rojas, Neural Networks - A Systematic Introduction. Available online at: <http://page.mi.fu-berlin.de/rojas/neural/>
- S. Theodoridis, K. Koutroumbas, Pattern Recognition, Academic Press.
- Students will be directed to papers available online and other Web resources on the subject.

5.13 Principles of Programming II (PoP II)

Teaching Staff

Keith Mannock

Online material

On Moodle.

Aims, Outline and Learning Outcomes

This module covers object-oriented programming, including the use of subclasses, modules, and library classes to create well-organised programs. A further aim is to enhance students understanding of making appropriate choices on the selection of algorithms, their implementation together with the required data structures (e.g. arrays, lists, trees, graphs, depth- and breadth-first search algorithms). To enable students to develop programs for modern multi-core architectures utilising functional programming constructs.

This module further develops the techniques described in the Principles of Programming I module. This module discusses issues specifically related to developing programs for large programming projects and for modern computer hardware architectures.

On successful completion of this module a student will be expected to be able to:

- Explain and exploit subclasses, inheritance and interfaces to produce modular, well-organised code.
- To utilise the relevant features of a programming language.
- Implement simple algorithms and data structures, both in sequential and parallel environments.
- Show knowledge of basic concepts and principles of object-orientation such as objects and classes, encapsulation, object state, coupling, cohesion and modularity.
- Provide evidence of the understanding of functional programming constructs.
- Write code that exploits the networking facilities of a modern programming language.

Syllabus

- Transition to Object-Oriented, including types, inheritance, etc.
- Static typing and reference types
- Further recursion, memorisation, etc.
- Local I/O
- Networking
- Functional programming constructs: Filter, map, reduce, etc.
- Style rules
- Further test driven development
- Avoid boilerplate
- Programming in teams

Prerequisites

Principles of Programming I module.

Assessment

This module is assessed by

- Two-hour unseen examination online or paper (80%)
- Programming assignments and group work (20%)

Reading

A set of course notes will be provided but the following text covers similar topics:

- Core Java for the Impatient by Cay Horstmann. Addison-Wesley. 2014. ISBN 978-0321996329

5.14 Semantic Technologies (ST)

Teaching Staff

Michael Zakharyashev

Online material

<http://www.dcs.bbk.ac.uk/~michael/st/st.html>

Aims

This module is a gentle introduction to Semantic Technologies that provide easier ways to find, share, reuse and combine information. Semantic Technologies define and link data on the web or within an enterprise by developing languages to express rich, self-describing interrelations of data in a form that machines can process. They provide an abstraction layer above existing IT technologies that connects data, content and processes. Semantic Technology standards developed by W3C include

- a flexible data model RDF (Resource Description Framework) for storing data in graph databases;
- schema and ontology languages for describing concepts and relationships (RDFS and OWL);
- the query language SPARQL designed to query data across various systems and databases and to retrieve and process data stored in RDF format.

Applications of Semantic Technologies range from Linked Data, Wikidata, Health-care and Pharma Industry, Supply Chain Management, Publishing and Media Management, Web Search and E-commerce to Data Integration in the Oil & Gas industry.

The aims of this module are to

- introduce the theoretical foundations of Semantic Technologies, including the languages RDF/S, SPARQL, the Web Ontology Language OWL;
- provide the students with practical skills of modelling data using RDF/S, querying RDF triplestores, and building ontologies;
- overview the current applications of Semantic Technologies in health care, media management, and industry;
- demonstrate a few standard algorithms for classification of concepts in ontologies.

Syllabus

- Introduction to the module. Ontologies in (Computer) Science. Knowledge graphs. Schema.org. Wikidata. Lab: building a Don Corleone family ontology.
- Is XML a semantic technology? The tree model of XML documents, XML Schema. Querying XML documents, XPath, JSON. Lab: building a pizza ontology.
- Resource Description Framework (RDF). RDF Schema. RDF/S semantics. Terse RDF Triple Language Turtle. Linked Data. Lab: extracting RDF data from natural language texts.
- SPARQL Query Language. Querying RDF triplestores. Lab: setting up and querying Apache Jena triplestore.
- Ontology-based data access (OBDA). OBDA platform Ontop. Lab: setting up ontology-based access to the IMDB database.
- Requirements for ontology languages. From RDFS to OWL. OWL ontologies.
- Ontology engineering. OWL ontologies in life sciences and industry. Lab: designing a travel agent's ontology.
- Open vs closed worlds. Reasoning with OWL. Introduction to Description Logic and formal semantics.

Prerequisites

None

Assessment

By 2-hour written examination and coursework exercises, weighting 80% and 20% respectively.

Reading

- G. Antoniou and F. van Harmelen. A Semantic Web Primer. MIT Press, 2004.
- P. Hitzler, M. Kroetzsch and S. Rudolph. Foundations of Semantic Web Technologies. Chapman & Hall, 2009.

5.15 Advances in Data Management (ADM)

Teaching Staff

Alex Poulouvassilis

Online material

<http://moodle.bbk.ac.uk/>

Aims

To study advanced aspects of database management and recent advances in data management technologies in three major directions: performance, distribution of data, and heterogeneity of data.

The module examines the technologies underlying modern database management systems. It studies advanced aspects of query processing, transaction management, distributed data management, and recent developments in web data, “big data” and alternative database architectures.

Syllabus

- Review of the fundamental principles of modern database management systems, relational databases and SQL.
- Query processing and query optimisation.
- Transaction management: ACID properties, concurrency control, recovery.
- Beyond records and objects: stored procedures and functions, triggers, semantic technologies.
- Distributed databases: data fragmentation and replication, distributed query processing, distributed transaction management.
- Heterogeneous data integration.
- XML data management.
- Linked Open Data.
- Parallel databases.
- Big data and NoSQL/NewSQL stores.
- Graph databases.

Prerequisites

A first module in Database Systems (e.g. as taught in a typical UK undergraduate degree in computer science) or the Birkbeck module Data and Knowledge Management.

Assessment

By 2-hour written examination and practical coursework, with weighting 90% and 10%, respectively.

Reading

- R. Ramakrishnan and J.Gehrke, Database Management Systems, McGraw-Hill 2003 (3rd Edition)
- A.Silberschatz, H.F.Korth and S.Sudarshan, Database System Concepts, McGraw-Hill 2011 (6th Edition)
- M. T. Oszu, P. Valduriez, Principles of Distributed Database Systems, Prentice-Hall 2011 (3rd Edition)
- A. Elmagarmid, M. Rusinkiewicz, A. Sheth (eds), Management of Heterogeneous and Autonomous Database Systems, Morgan Kaufmann, 1999
- Research papers will be distributed to students; students will also be directed to Web resources on the subject.

6 MSc Project

Please refer to the page

http://www.dcs.bbk.ac.uk/dcswiki/index.php/MSc_CS_project

for a detailed description and guidelines.

Aims

In the MSc project students will be able to demonstrate their skills in organising and completing a task that goes beyond a typical coursework assignment. This includes planning and executing a major piece of programming work or an experimental campaign appropriate to the MSc programme and presenting existing approaches in the problem area (placing the student's own approach in the wider technical and conceptual context).

The MSc project will offer students the opportunity to:

- develop a systematic understanding and critical awareness of an agreed problem relevant to Data Science, as described in a project proposal form
- plan and execute a major piece of programming work appropriate to the MSc programme
- critically present existing approaches in the problem area, place their own approach in the wider area and evaluate their contribution
- gain experience in communicating complex ideas and concepts to others by writing a comprehensive, self-contained report.

Students are encouraged to come up with their own ideas for projects. In order to arrange supervision for the project a student should discuss possible projects with the Projects Tutor, Programme Director or with the lecturer who seems the most appropriate for the topic.

Syllabus

The main part of the module will be done by the students on their own (supported by the supervisor). There is a small taught part of the module in which the students are acquainted with

- how to formulate the objectives/aims of an MSc project
- how to write a project proposal
- how to organise and plan the project
- how to research literature
- how to write a project report.

Assessment

Written project proposal (of about 2000-3000 words) and written project report (of about 10,000 words), weighting 20% and 80%, respectively.

Reading

- As recommended by the supervisor.

7 Administration and Assessment

For detailed College rules and regulations see

<http://www.bbk.ac.uk/mybirkbeck/services/rules>

and, in particular,

<http://www.bbk.ac.uk/mybirkbeck/services/rules/casregs.pdf>

Below we summarise the most relevant rules for the MSc Data Science.

7.1 Requirements for the Award of the MSc

Each taught module (all modules except the project) is assessed by a written exam and, in most cases, by additional coursework. The project module is assessed by the project proposal document (20%) and the project report (80%).

For each module, a Pass requires at least 50% of the available marks (computed according to the corresponding weights of the parts of the assessment). Up to 30 credits of the taught modules with a mark between 40% and 49% can be compensated (assuming that the total weighted average mark is above 50%) on the MSc (note that College regulations do not allow compensation on PGDip and PGCert). Additionally, there is a 60-credit project module.

To gain an award, the following is required:

- Postgraduate Certificate (PGCert): pass 4 compulsory 15-credit modules.
- Postgraduate Diploma (PGDip): pass all compulsory taught modules and one optional taught module.
- Master of Science (MSc): requirements for PGDip and pass the project.

The final grade is computed by taking the weighted average (according to the credits) of the module assessment marks. The following has to be satisfied:

- Pass requires at least a 50% weighted average pass mark
- Merit requires at least a 60% weighted average pass mark
- Distinction requires at least a 70% weighted average pass mark and, normally, at least 70% on the project.

7.2 Announcement of Results

The Examination Board meets in July to consider the results of the written exams and coursework, and in November to consider the results of the projects and to award degree.

Shortly after the meeting of the exam board you will receive a letter from the Department about your results. Your results and grades will be officially confirmed by the College on your MyBirkbeck profile.

Students who have not paid their fees won't be served any information about their examination results.

7.3 Choosing the Optional Modules

You will receive a form from the Programme Administrator at the beginning of the relevant term to indicate your choice of the optional modules. You have to return this form by the specified deadline so that your chosen module can be listed among the modules that you wanted to be assessed that year.

Optional module availability is subject to timetabling constraints and student demand. In the event that an optional module is over-subscribed, available places will be allocated on a first-come, first-served basis determined by the date you return your module choice form to the Programme Administrator.

7.4 Mitigating Circumstances and Deferral

A Mitigating Circumstances claim should be submitted if valid detrimental circumstances result in:

- the late or non-submission of assessment;
- non-attendance of examination;
- poor performance in assessment.

If a student feels their circumstances warrant consideration by the Board of Examiners they should notify the Programme Director, in writing, in advance, at the earliest opportunity (within 7 days of the assessment deadline or examination) using a *Mitigating Circumstances Claim Form*, which can be downloaded from MyBirkbeck or from:

<http://www.bbk.ac.uk/registry/policies/documents/MitCircs.pdf>

In the form, students should state whether the circumstances relate to non-attendance at an examination or late submission of an assignment and should include supporting evidence (e.g. a medical certificate giving the nature and duration of any illness). Students should be aware that discussing their claim with a member of staff does not constitute a submission of a claim of mitigating circumstances.

For a claim to be accepted a student must produce independent documentary evidence to show that the circumstances:

- have detrimentally affected their performance/submission/attendance in assessment or will do so;
- were unforeseen;
- were out of their control and could not have been prevented;
- relate directly to the timing of the assessment affected.

For further information, students may consult the document on mitigating circumstances through MyBirkbeck:

<http://www.bbk.ac.uk/reg/regs/mitcircspol>

In **exceptional cases**, students may be permitted to defer the written exams and/or the project to the following year. They must apply by filling in a *Mitigating Circumstances Claim Form*. Students have to do this before **May 1st** for exams and by **September 1st** for the project.

A student who defers an element of assessment normally has to enter for that element the following year; usually no further deferrals are permitted. Simply not turning up for an exam or failing to submit a coursework or project, without permission to defer, will be considered to be the same as failing it, in the sense that it will count as one of the two attempts that you are permitted to make at passing that element.

7.5 Retake and Reassessment

One reassessment (but only one) is allowed for each element. You may be reassessed in a failed coursework, written exam or the project if your marks for that module are below 50%. If your marks are below 40%, then you have to retake the whole module (i.e., attend lectures and be reassessed in each element of the module, including coursework and exam).

Students who fail an assessment and awarded a reassessment opportunity have their reassessment subject to a cap of 50% for the reassessed element. The cap does

not apply to a retake of a whole module and to students with accepted mitigating circumstances.

There are no special resit exams; students resit alongside the other candidates in May/June the following year. They normally do so a year after their first attempt. Where the syllabus has changed, we set a paper that is suitable for resit candidates, providing alternative questions where necessary. Note, however, that we do this only for candidates from the previous year, not from further in the past.

Also note that part-time students normally need to accumulate at least 45 credits (out of the available 75) in their first year in order to progress into the second year. Students who do not achieve this will not be able to complete their studies in two years. Instead they will have to spend at least one year as a repeating student, retaking the failed modules.

Under normal circumstances this would take place the following year and students would not be allowed to take any new modules until they had passed the failed modules. However, because some modules on this programme are taught in the evenings on alternate years only, doing so would mean that students in their third year could have no new modules available to them in the evenings. This would have the effect of extending the duration of the degree to at least four years.

As a result, we permit students in such circumstances to enrol on four new modules in their second year, postponing their second attempts at the failed modules to the third year. This is not ideal, but seems preferable to extending the duration of the degree further.

7.6 Re-enrolment

Repeat students, i.e., students who have to retake some modules (and are not taking any new modules) will be charged pro-rata based on the number of credits they retake.

Assessment only students, i.e. those students who

- are being reassessed for coursework and/or examinations only
- have deferred their examinations and are not taking any new modules
- have deferred the project and do not require supervision (resubmitting only)

pay a reduced fee that will allow them access to College facilities (Library and workstation rooms). While deferred students are classed as assessment only they are allowed to attend lectures for revision purposes. They should formally seek the permission of module tutors to ensure classes are not oversubscribed.

Dissertation only students, i.e. students who retake the project with supervision, pay one third of full fees. Note that

- a student who has to resubmit the dissertation and be reassessed for examination or coursework will be progressed as dissertation only
- a student who has to resubmit the dissertation and also repeat modules will be progressed as repeat and fees are based pro-rata on the number of credits.

7.7 Examinations

Please consult the programme's intranet web page (for enrolled students):

http://www.dcs.bbk.ac.uk/dcswiki/index.php/MSc_DS_Intranet

7.8 Coursework

A number of modules require students to submit coursework as part of the assessment. Please consult the web page of the relevant module or contact the teaching staff of the module for particular details.

Submitted coursework must always be the students own work, except where explicitly noted. Students are required to confirm in writing or via e-mail that each item of coursework submitted is indeed their own work.

The Department and College have strict guidelines and penalties associated with plagiarism, and routinely submit students' work to plagiarism detection services. More details are given in the section "Assessment Offences and Plagiarism" of this booklet.

College policy dictates how Schools will treat work that is due for assessment but is submitted after the published deadline. Any work that is submitted for formal assessment after the published deadline but before the cut-off date (normally ten working days after the deadline) is given two marks: a penalty mark of 50% for postgraduate students, assuming it is of a pass standard, and the real mark that would have been awarded if the work had not been late.

Both marks are given to the student on a cover sheet. If the work is not of a pass standard a single mark is given.

If you submit late work that is to be considered for assessment, then you should submit a mitigating circumstances form, see above, and provide written documentation, medical or otherwise, to explain why the work was submitted late. The case will then be considered by the appropriate sub-board or delegated panel of the Board of Examiners.

If no case is made then the penalty mark will stand. If the case is made and accepted then the examination board may allow the real mark to stand.

7.9 Projects

Please consult the programme's intranet web page (for enrolled students):

http://www.dcs.bbk.ac.uk/dcswiki/index.php/MSc_DS_Intranet

7.10 Assessment Offences and Plagiarism

Please see at MyBirkbeck

<http://www.bbk.ac.uk/student-services/exams/assessment-offences>

for the College Policy on Assessment Offences.

One particular assessment offence is *plagiarism* that is defined as

“[...] copying a whole or substantial parts of a paper from a source text (e.g., a web site, journal article, book or encyclopedia), without proper acknowledgement; paraphrasing of another's piece of work closely, with minor changes but with the essential meaning, form and/or progression of ideas maintained; piecing together sections of the work of others into a new whole; procuring a paper from a company or essay bank (including Internet sites); submitting another student's work, with or without that student's knowledge; submitting a paper written by someone else (e.g., a peer or relative), and passing it off as one's own; representing a piece of joint or group work as one's own.”

Also,

“[a] student who knowingly assists another student to plagiarise (for example by willingly giving them their own work to copy from) is committing an examination offence.”

The College considers plagiarism a serious offence, and as such it warrants disciplinary action. This is particularly important in assessed pieces of work where plagiarism goes so far as to dishonestly claim credit for ideas that have been taken by someone else.

The College also provides learning support for exams and assessments, please see

<http://www.bbk.ac.uk/student-services/learning-development>

and guidelines on plagiarism

<http://www.bbk.ac.uk/student-services/exams/plagiarism-guidelines>

8 Student Services

The College provides various services to students, see:

<http://www.bbk.ac.uk/student-services>

In particular, there are the Counselling Service, the Disability and Dyslexia Service, and the Mental Health Service. They provide specialist support to students.

8.1 Counselling Service

The Counselling Service

<http://www.bbk.ac.uk/student-services/counselling-service>

provides assistance to students who are experiencing emotional difficulties which may be impacting upon their studies or overall experience at Birkbeck.

8.2 Mental Health Service

Many students experience mental health difficulties at some point in their time at university. Whether you have a formally diagnosed psychiatric condition or other form of mental health difficulty such as anxiety or depression, we encourage you to seek support in your studies. Birkbecks Mental Health Service

<http://www.bbk.ac.uk/student-services/mental-health-advisory-service>

is a first point of contact for students experiencing mental health issues at any stage during their studies.

8.3 Disability and Dyslexia Service

At Birkbeck we welcome students with disabilities. We aim to provide all of our students with a study environment that enables them to participate fully in our courses.

The Disability and Dyslexia Service:

<http://www.bbk.ac.uk/student-services/disability-service>

can provide advice and support to students with conditions that impact their ability to study, such as:

- specific learning difficulties (dyslexia, dyspraxia, dyscalculia, AD(H)D)

- sensory impairments (blind/partially sighted, deaf/hearing impaired)
- mobility conditions (including RSI, arthritis, neck back and knee conditions etc.)
- medical conditions (e.g. HIV, CFS, diabetes, cancer, chest and respiratory conditions etc.)
- autism-spectrum conditions (autism or Aspergers syndrome)

They can provide support during your studies including

- Your Study Support Plan
- The Disabled Students' Allowance
- Access to Learning Fund
- Charities and trusts
- Dyslexia screening test
- Government benefits
- Personal emergency evacuation plans
- Pager alert system
- Rest Room
- Toilet facilities
- Car parking
- Disability and Dyslexia Support in the Library and IT Services

9 The Business Engagement Team

The School of Business, Economics and Informatics has a dedicated Business Engagement team to provide you with extra support. The team delivers the following initiatives to support you in your career aspirations:

9.1 Mentoring Pathways

Mentoring Pathways pairs successful applicants with industry professionals for individual advice and guidance. There are a number of places available for final year undergraduates and postgraduate students. We have partnerships with a number of key organisations and work alongside Birkbeck alumni who provide mentors. Applications open in the autumn. Please email mentoring@bbk.ac.uk or visit

<http://www.bbk.ac.uk/business/business-services/mentoring-external>

for more information.

9.2 Enterprise Pathways

Whether you are setting out in your journey as an entrepreneur or have already established a thriving business, we offer various pathways to support you. These include a non-credit bearing module with workshops once a month throughout the academic year, access to digital resources, enterprise boot camps and inspirational talks to help you to develop your ideas and network with other students. Please email enterprise@bbk.ac.uk or visit www.bbk.ac.uk/enterprise.

9.3 Keeping in Touch

You can also follow BEI on social media for information and conversations:

- Twitter: @BirkbeckBEI
- Facebook, Google+ and LinkedIn: Search 'BirkbeckBEI'

Please visit our website

www.bbk.ac.uk/business/business-services

for more resources and information.

10 Career Development

Most students are interested in developing their careers, either within their current field of work or in a completely new direction. The Careers Group, University of London

<http://www.thecareersgroup.co.uk/>

offers great expertise and experience in working with students and graduates of all ages and at all stages of career development.

The Careers and Employability Service

<http://www.bbk.ac.uk/careers/careers-service>

is our in-house service for enhancing career development and employability throughout your time at Birkbeck, from enrolment through to graduation.

There is also Birkbeck Talent, a professional recruitment service aimed exclusively at assisting Birkbeck students to find work whilst studying and after graduation. They work with London's top employers to offer innovative internships, prestigious job vacancies and exciting graduate opportunities. To find out more please visit

<http://www.bbk.ac.uk/student-services/birkbeck-talent-service>

