

MSc Data Science

Programme Specification

Awarding Institution:

University of London (Interim Exit Awards made by Goldsmiths' College)

Teaching Institution: Goldsmiths, University of London

Name of Final Award and Programme Title: MSc Data Science

Name of Interim Exit Award(s): Postgraduate Diploma in Data Science

Duration of Programme: 1 year full-time or 2 years part-time

UCAS Code(s): Not applicable

HECoS Code(s): (100366) Computer Science

QAA Benchmark Group: Computing

FHEQ Level of Award: Level 7

Programme accredited by: Not applicable

Date Programme Specification last updated/approved: December 2020

Home Department: Computing

Department(s) which will also be involved in teaching part of the programme:

Not applicable

Programme overview

Computer Science is one of the main research areas in the Department of Computing at Goldsmiths, and is a key component of the multiple interdisciplinary research directions, in particular of Data Science, which are developed within the Department. Both the science of efficient computing and the science of extracting information from data continue to increase in importance in various disciplines in which the large volume and the complexity of data used in research impose unprecedented challenges to the data analysis approaches traditionally employed in these disciplines. Tractable and efficient solutions are needed for big data management and analysis nowadays, and Data Science is the answer.

The Data Science MSc programme at Goldsmiths, which is built upon the existing expertise in Statistical Computing, Data Mining, Machine Learning, Time Series Forecasting, Soft Computing, Big Data, Algorithmics and Databases in the Computing Department, responds to the increasing need for applying modern computing approaches to data processing and analysis in new monodisciplinary and interdisciplinary research in the Department and the College, and for providing students with postgraduate training in this area.

The programme is well integrated in the multidisciplinary teaching and research activity in the College. In research terms, the Data Science MSc programme will complement the research group in Data Science and Soft Computing that students can join for their final project or for PhD studies subsequently. Students will equally have the opportunity to join other research teams in the College for their final project work, in which they can contribute with their gained expertise in Data Science to the interdisciplinary work developed by those teams.

In particular students will be exposed to such research work in Data Science or to multidisciplinary research work employing Data Science via the Data Science Research Topics 15-credit module in which researchers from across the College will present their work and attract students' interest to possibly contribute to that work during a final project. In the same module students will have the opportunity to attend presentations by professionals specialised in Data Science, based on current common concerns and cutting-edge specific technologies and applications used in industry.

Students will receive highly specialised training in various conventional and modern statistical, machine learning and data mining algorithms for data analytics, and in highly scalable methods and technologies used in the management and analysis of big data. These topics form the basis of the two compulsory 30-credit modules, namely Machine Learning and Statistical Data Mining on one hand, and Big Data Applications on the other. In particular students will learn these concepts and techniques through the use of major software technologies for data analysis and mining, including text and web mining, such as R, IBM SPSS Statistics and RapidMiner, and on big data software technologies as Hadoop based on HDFS distributed file system and the MapReduce scalable computing approach, NoSQL databases, Hive data warehousing, Pig Latin, as well as general-purpose programming languages as Python. These software technologies will be running on a dedicated computer cluster sustaining large applications.

Students will be exposed to real world applications by modelling and implementing applications encountered in business (including customer analytics, credit scoring, financial forecasting, reality mining, mobile telephone data analysis), in health and medical research (including automatic diagnosing, genetic data mining and bioinformatics, mining online medical publication libraries), in unstructured data analysis (including text and web mining applied in sentiment analysis and the intelligent web). Further specialised training is provided in elective modules which are shared with other programmes. These electives include Neural Networks, Introduction to Natural Language Processing, Linked Data and the Semantic Web, Interaction Design, Artificial Intelligence, Open Data, and Music Information Retrieval (subject to timetabling and staff availability). These modules provide further specialisation in various topics related to forecasting and soft computing, web mining, information retrieval, intelligent systems, and open data, which are directly linked or complement the Data Science topics studied in the compulsory modules of the programme.

The Final Project module, worth 60 credits, will allow students to undertake substantial independent projects that will allow them to demonstrate project planning and management skills, research skills, and written and oral presentation skills. Students will integrate their knowledge and will use skills acquired in the other programme's modules in the implementation of their final project in Data Science or related interdisciplinary topics.

Programme entry requirements

Upper second class undergraduate degree or above in computing, mathematical sciences, physics, engineering, finance, bioinformatics and computational biology, social sciences with a good basis of statistical knowledge and some programming experience.

Prospective students should have an interest in one or more Data Science related topics including statistics and data analysis, machine learning, data mining, big data, bioinformatics, intelligent web, financial forecasting, and computational social sciences.

Outstanding practitioners or individuals with strong commercial experience may be considered. Nonnative English students should normally have a minimum IELTS score of 6.5 or equivalent.

Aims of the programme

The aim of this Programme is to produce graduates who are autonomous, creative and reflective computing practitioners, and have in particular practical skills and research abilities in the Data Science field. Our graduates should have:

- Knowledge of computing technologies across a range of specialist topics in Data Science, both in terms of the latest research advances and industry standards.
- The ability to design and implement from small to large scale data analysis processes, by using existing specialised software solutions or by implementing their own software solutions, on usual or specialised scalable computing hardware, including computing clusters.
- Strong transferable skills, in particular the ability to work independently and in groups and reflectively evaluate their own work.
- The ability to perform monodisciplinary research in Data Science and interdisciplinary research in work involving also data analysis or other topics related to Data Science, or the ability to work with industry. In particular our students are encouraged to join a research team in the College, or undertake industry internships as part of their final project.

What you will be expected to achieve

Learning outcomes for the Postgraduate Diploma:

Knowledge and understanding

Code	Learning outcome	Taught by the following module(s)
A1	A deep and practical understanding of state of the art of scalable computing technology particularly applied on variable volumes of data	Data Programming, Big Data Applications, Machine Learning and Statistical Data Mining, Data Science Research Topics, Linked Data and the Semantic Web, Open Data
A2	A systematic understanding of machine learning and statistical data mining techniques used in data analytics and in other related areas	Machine Learning and Statistical Data Mining, Data Science Research Topics, Neural Networks, Introduction to Natural Language Processing, Music Information Retrieval
A3	A good understanding of various topics involved in interdisciplinary research in which Data Science is applied	Data Science Research Topics, Machine Learning and Statistical Data Mining, Final Project, Interaction Design, Music Information Retrieval
A4	A critical awareness of practical and theoretical contexts in which data scientists work	This will be taught throughout the programme
A5	Critically analyse the application of technology to real world problems particularly in industry and interdisciplinary research	This will be taught throughout the programme

Cognitive and thinking skills

Code	Learning outcome	Taught by the following module(s)
B1	Apply advanced skills and research-lead specialist knowledge in the design of software and data analyses; academic writing and presentation skills	This will be taught throughout the programme
B2	Critically analyse the application of technology to real world problems particularly in industry and interdisciplinary research	This will be taught throughout the programme

Subject specific skills and professional behaviours and attitudes

Code	Learning outcome	Taught by the following module(s)
C1	Apply advanced skills and research-lead specialist knowledge in the design of software and data analyses; academic writing and presentation skills	This will be taught throughout the programme
C2	Critically analyse the application of technology to real world problems particularly in industry and interdisciplinary research	This will be taught throughout the programme

Transferable skills

Code	Learning outcome	Taught by the following module(s)
D1	Be able to do academic research and writing, and present themselves and their work	Data Science Research Topics, Final Project
D2	Be able to reflect on and evaluate their work	This will be taught throughout the programme
D3	Be able to work effectively in groups	This will be taught in group projects which will be part of the assessment in modules including Machine Learning and Statistical Data Mining, and Big Data Applications
D4	Be able to work effectively in groups	All taught modules

Additional learning outcomes for the MSc:

Knowledge and understanding

Code	Learning outcome	Taught by the following module(s)
A1	Applied a deep understanding of cutting edge technologies in the creation of a substantial commercially and/or research-wise relevant project	Final Project

Cognitive and thinking skills

Code	Learning outcome	Taught by the following module(s)
B1	Propose, plan, execute and evaluate a significant piece of original work	Final Project

Subject specific skills and professional behaviours and attitudes

Code	Learning outcome	Taught by the following module(s)
C1	Independently and in cooperation research and apply state of the art technologies in the context of concrete problems related to scalable data management and analytics	Final Project, but also the other modules in various degrees, in particular Machine Learning and Statistical Data Mining, and Big Data Applications
C2	Design and program advanced computer software and data products	Final Project, Machine Learning and Statistical Data Mining, and Big Data Applications, Introduction to Natural Language Processing

Transferable skills

Code	Learning outcome	Taught by the following module(s)
D1	Be able to do academic research and writing, and present themselves and their work	Data Science Research Topics, Final Project
D2	Be able to reflect on and evaluate their work	This will be taught throughout the programme
D3	Be able to work effectively in groups	This will be taught in group projects which will be part of the assessment in modules including Machine Learning and Statistical Data Mining, and Big Data Applications
D4	Be proactive, plan their activity in advance, and exercise personal responsibility in their work	This will be taught in throughout the programme

How you will learn

The Department of Computing is committed to a diverse and stimulating range of learning and teaching methods that ensure the programme outcomes are addressed rigorously and

effectively. Learning emphasises a close synthesis between theoretical understanding and practical application that helps students develop an advanced, critical approach to the subject of Computing in general and to Data Science in particular.

The various modules of the programme provide a diverse range of topics. These will be further developed through students' independent research and learning activities directed towards module assignments and the large-scale project component. The department is committed to providing a diverse and innovative range of teaching styles across its degree programmes. These include traditional lecture and laboratory sessions but also a range of more interactive and self directed activities focusing on independent, creative work and self presentation. The nature of the learning activities will vary between different modules, but includes data analysis, design and implementation (concerning several sorts of applications in industry and/or interdisciplinary research), software design and coding, devising data products, project planning, group activity and creative and research work.

Students are expected to engage in considerable independent reading and practical work for all modules culminating in the final year project. This independent work will be supported by library resources, access to lab space and computing cluster facilities, and supervision from teaching staff.

How you will be assessed

The Department is committed to providing diverse types of assessment. Our methods of assessment are designed to reflect research and business relevant activities and to encourage independent and collaborative work. In particular our assessment integrates different kinds of hands-on practical work including software and data products development, data mining/analysis design and implementation, business planning, and individual or group work. Students will be required to present their work in a number of different ways including posters, traditional reports and short management reports, oral presentations, and software and data products live or recorded presentations.

Each module in the programme will have its own type of assignment which corresponds to the nature of the module. In addition to usual assignments, students will have a major final project in the summer term, which should integrate what students have learnt throughout the programme. The final project is an opportunity for students to work independently in a large project reflecting the state of the art technology at a research level.

In collaborative assignment work particular care must be taken by students to describe clearly and in detail precisely the nature of their contributions, and the contributions of their group collaborators. This must be delivered as part of students' assignment written reports and/or evaluation documentation, as required in the documentation brief. In particular

students should report on the group dynamics, explain their role in the group, and indicate any problems that the group encountered and how they have been addressed.

Feedback is very important to the learning process, and shows students how to improve their work, and provides suggestions on how to learn more effectively in the future.

Therefore the Department is committed to providing timely and full feedback on all assessed assignments.

Final projects will be assessed based on the submission of a final report and a presentation in a viva. Guidance on the structure and writing of the report will be given in the module handbook. Moreover, general guidance on writing scientific work will specifically be provided in the Data Science Research Topics module. Projects will be marked by a panel composed of two members of academic staff.

Students who are unable to submit an assessment on time due to illness or other unavoidable circumstances, must provide documentary evidence to their personal tutor in order to be allowed a late submission. Evidence must also be supplied for students to apply for consideration of mitigating circumstances in assessment.

Marking criteria

Mark	Descriptor	Specific Marking Criteria
80-100%	Distinction (Outstanding/ Exceptional)	A grade in the range of 80-100% will be awarded in the case of really accomplished work that demonstrates high levels of scholarship and originality. This grade will reflect the overall achievement of the appropriate learning outcomes to an exceptionally accomplished level. In particular a grade in the 90s should be reserved for work deemed to be outstanding, and of publishable quality.
70-79%	Distinction	A grade in the range of 70-79% will be awarded when candidates show evidence of an excellent application of appropriate knowledge, understanding and skills as specified in the module learning outcomes. Demonstration of a thorough grasp of relevant concepts, methodology and content appropriate to the subject discipline; indication of originality in application of ideas, in synthesis of material or in performance; insight reflects depth and confidence of understanding of the material.
60-69%	Merit	Demonstration of a deep level of understanding based on a competent grasp of relevant concepts, methodology and content; display of skill in applying interpreting complex

Mark	Descriptor	Specific Marking Criteria
		material; organisation of material at a high level of competence. Students should be able to demonstrate the ability to work independently to research and implement state of the art technologies.
50-59%	Pass	Demonstration of a sound level of understanding based on a competent grasp of relevant concepts, methodology and content; display of skill in organising, discussing and applying complex material. Students should be able to implement state of the art technologies under guidance.
30-49%	Fail	Represents an overall failure to achieve the appropriate learning outcomes. Students achieve some of the aims but were unable to demonstrate independence and originality beyond what would be expected at undergraduate level.
10-29%	Bad fail	Represents a significant overall failure to achieve the appropriate learning outcomes.
1-9%	Very bad fail	A submission that does not attempt to address the modules specified learning outcomes. It will be considered a non-valid attempt and the module must be re-sat.
0%	Non submission or plagiarised	Work was not submitted, or it was plagiarised.

How the programme is structured

MSc Data Science consists of two terms of taught modules followed by a large-scale project. The taught modules introduce you to fundamental mathematical and computational skills and show you how to apply them to real world data.

The programme includes:

- A firm grounding in the theory of data mining, statistics and machine learning
- Hands-on practical real world applications such as social media, biomedical data and financial data with Hadoop (used by Yahoo!, Facebook, Google, Twitter, LinkedIn, IBM, Amazon, and many others), R and other specialised software
- The opportunity to work with real-world software such as Apache

In addition to the compulsory programme modules, students choose 3 optional modules from the indicative list (revised each year) below.

Full-time mode

Module Title	Module Code	Credits	Level	Module Status	Term
Mathematics and Statistics for Data Science	IS71104A	15	7	Compulsory	1
Data Programming	IS71068A	15	7	Compulsory	1
Machine Learning	IS71071A	15	7	Compulsory	2
Big Data Applications	IS71059B	15	7	Compulsory	2
Data Science Research Topics	IS71058A	15	7	Compulsory	2
Three 15 credit optional modules totalling 45 credits from the list below:	Various	45	7	Optional	1 and 2
Neural Networks	IS71040A	15	7	Optional	1
Open Data	IS71063A	15	7	Optional	1
Data Visualisation and the Web	IS71082A	15	7	Optional	
Data Compression	IS71067A	15	7	Optional	1
Multimedia Informatics I: Music Information Retrieval and digital musicology	IS71064A	15	7	Optional	2
Geometric Data Analysis	IS71069B	15	7	Optional	2
Final Project	IS71061A	60	7	Compulsory	3

Part-time mode

Academic year of study 1

Module Title	Module Code	Credits	Level	Module Status	Term
Mathematics and Statistics for Data Science	IS71104A	15	7	Compulsory	1
Data Programming	IS71068A	15	7	Compulsory	1
Machine Learning	IS71071A	15	7	Compulsory	2
Big Data Applications	IS71059B	15	7	Compulsory	2

Academic year of study 2

Module Title	Module Code	Credits	Level	Module Status	Term
Final Project	IS71061A	60	7	Compulsory	3

Module Title	Module Code	Credits	Level	Module Status	Term
Data Science Research Topics	IS71058A	15	7	Compulsory	2
PLUS 3 X 15 CAT optional (indicative) modules from the FT route above	Various	45	7	Optional	1 or 2

Academic support

Support for learning and wellbeing is provided in a number of ways by departments and College support services who work collaboratively to ensure students get the right help to reach their best potential both academically and personally.

All students are allocated a Personal Tutor (one in each department for joint programmes) who has overall responsibility for their individual progress and welfare. Personal Tutors meet with their student at least twice a year either face-to-face, as part of a group and/or electronically. The first meeting normally takes place within the first few weeks of the autumn term. Personal Tutors are also available to students throughout the year of study. These meetings aim to discuss progress on modules, discussion of the academic discipline and reports from previous years if available (for continuing students). This provides an opportunity for progress, attendance and assessment marks to be reviewed and an informed discussion to take place about how to strengthen individual learning and success.

All students are also allocated a Senior Tutor to enable them to speak to an experienced academic member of staff about any issues which are negatively impacting their academic study and which are beyond the normal scope of issues handled by Programme Convenors and Personal Tutors.

Students are provided with information about learning resources, the [Library](#) and information available on [Learn.gold \(VLE\)](#) so that they have access to department/ programme handbooks, programme information and support related information and guidance.

Taught sessions and lectures provide overviews of themes, which students are encouraged to complement with intensive reading for presentation and discussion with peers at seminars. Assessments build on lectures and seminars so students are expected to attend all taught sessions to build knowledge and their own understanding of their chosen discipline.

All assessed work is accompanied by some form of feedback to ensure that students' work is on the right track. It may come in a variety of forms ranging from written comments on a

marked essay to oral and written feedback on developing projects and practice as they attend workshops.

Students may be referred to specialist student services by department staff or they may access support services independently. Information about support services is provided on the [Goldsmiths website](#) and for new students through new starter information and induction/Welcome Week. Any support recommendations that are made are agreed with the student and communicated to the department so that adjustments to learning and teaching are able to be implemented at a department level and students can be reassured that arrangements are in place. Opportunities are provided for students to review their support arrangements should their circumstances change. The [Disability](#) and [Wellbeing](#) Services maintain caseloads of students and provide on-going support.

The [Careers Service](#) provides central support for skills enhancement, running [The Gold Award](#) scheme and other co-curricular activities that are accredited via the Higher Education Achievement Report ([HEAR](#)).

The [Academic Skills Centre](#) works with academic departments offering bespoke academic literacy sessions. It also provides a programme of academic skills workshops and one-to-one provision for students throughout the year.

Links with employers, placement opportunities and career prospects

The MSc Data Science develops analytic and critical skills, providing successful students with the tools and competencies needed to intelligently interrogate numerical, textual and qualitative data; to extract meaning from raw information; and to communicate the results of their investigations, and their implications, to stakeholders or other interested parties. These skills lead naturally to embarking on a variety of careers, with employers from the financial sector, technology firms small and large, biomedical research sector, the charitable and voluntary sector, and public research sector. The mix of attributes encouraged by the programme as technical ability to manage and process data, reflection and insight to develop understanding, and empathy and awareness to communicate it effectively is highly desirable to prospective employers. The programme's structure, in particular around the final project and preparation for it, encourages student engagement with projects requiring the services of data scientists, and will provide networking opportunities to start the students along their chosen career path. Our graduates will be exposed to the ethical issues of data science; in the modern era of data availability, it is crucial that all participants in data exchange and treatment are aware of the impact of their behaviour on privacy, anonymity and personal security.

The requirements of a Goldsmiths degree

All taught postgraduate degrees have a minimum total value of 180 credits and involve one calendar year of full-time study. Some programmes may extend over more than one calendar year and, when this is the case, they have a higher total credit value. Programmes are composed of individual modules, each of which has its own credit value. Part-time students normally take modules to the value of 90 credits each year. If a programme has a part-time pathway, the structure will be set out in the section “How the programme is structured” above. Normally, all modules are at level 7 of the Framework for Higher Education Qualifications.

More detailed information about the structure and requirements of a Goldsmiths degree is provided in the [Goldsmiths Qualifications and Credit Framework](#).

Modules

Modules are defined as:

- “Optional” – which can be chosen from a group of modules
- “Compulsory” – which must be taken as part of the degree

Progression

Some programmes may require students to pass specific modules prior to completion of the dissertation/major project (or equivalent). Additionally, where a programme of study extends beyond one calendar year, students may be required to pass specific modules in their first year of study before progressing to the second year. Where this is the case, these requirements will be set out in this Programme Specification.

Award of the degree

In order to graduate, students must successfully complete all modules specified for the programme, as set out within the section “How the programme is structured” above.

Classification

Final degree classification is calculated on the basis of a student’s mean average mark (based on credit value) across all modules on the programme.

Masters degrees are awarded with the following classifications:

Distinction – 70%+

Merit – 60-69%

Pass – 50-59%

More detail on the [calculation of the final classification](#) is on our website.

Interim exit awards

Some programmes incorporate interim exit points of Postgraduate Certificate and/or Postgraduate Diploma, which may be awarded on the successful completion of modules to the minimum value of 60 credits or 120 credits respectively. The awards are made without classification.

When these awards are incorporated within the programme, the relevant learning outcomes and module requirements will be set out within the “What you will be expected to achieve” section above.

The above information is intended as a guide, with more detailed information available in the [Goldsmiths Academic Manual](#).

Programme-specific rules and facts

In order to progress to the Final Project, students must fulfil the requirements for a pass at Postgraduate Diploma level (pass all 120 credits of taught modules).

General programme costs

In addition to your tuition fees, you will be responsible for meeting standard costs associated with your study. Find out more information at gold.ac.uk/programme-costs.

Specific programme costs

Not applicable.

How teaching quality will be monitored

Goldsmiths employs a number of methods to ensure and enhance the quality of learning and teaching on its programmes.

Programmes and modules are formally approved against national standards and are monitored throughout the year, such as in departmental committees, a variety of student feedback mechanisms and through the completion of module evaluation questionnaires. Every programme has at least one External Examiner who reviews comments annually on the standards of awards and student achievement. External Examiner(s) attend Boards of Examiners meetings and submit an annual written report.

Modules, programmes and/or departments are also subject to annual and periodic review internally, as well as periodic external scrutiny.

Quality assurance processes aim to ensure Goldsmiths' academic provision remains current, that the procedures to maintain the standards of the awards are working effectively and the quality of the learning opportunities and information provided to students and applicants is appropriate.

Detailed information on all these procedures are published on the [Quality Office web pages](#).