Programme Specification 2020-21

MRes in Environmental Data Science

<table>
<thead>
<tr>
<th>Awarding body</th>
<th>University of Cambridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teaching institution</td>
<td>University of Cambridge</td>
</tr>
<tr>
<td>Accreditation details</td>
<td>None</td>
</tr>
<tr>
<td>Name of final award</td>
<td>MRes</td>
</tr>
<tr>
<td>Programme title</td>
<td>MRes in Environmental Data Science</td>
</tr>
<tr>
<td>HECoS code(s)</td>
<td>100394 (earth sciences)</td>
</tr>
<tr>
<td>Relevant QAA benchmark statement(s)</td>
<td>None</td>
</tr>
<tr>
<td>Qualifications framework level</td>
<td>7</td>
</tr>
<tr>
<td>Date specification produced</td>
<td>March 2021</td>
</tr>
</tbody>
</table>

The MRes in Environmental Data Science is offered by the UKRI Centre for Doctoral Training in Applications of Artificial Intelligence (AI) to the Study of Environmental Risk (AI4ER). The MRes Environmental Data Science is an integral part of a 1+3 PhD training programme and cannot be applied for separately as a stand-alone one-year degree. The Department of Earth Science will act as host to this multidisciplinary course, with the management team also including members from Engineering, Chemistry and Computer Science and Technology, alongside the British Antarctic Survey. Additionally, we have commitments to supervise students from AI4ER by over 50 qualified PhD supervisors comprising academics from 13 University departments as well as researchers at BAS.

Educational aims

We have designed an MRes that will bring a potentially diverse group (student recruitment is expected to be principally from Engineering, Computer Science, Physical Sciences and Mathematics, but numerate and interested Biological Scientists may also be anticipated) together into a cohesive cohort, all capable of innovation in the "AI for the study of environmental risk" domain, skilled in the methods to be deployed in PhD-level research.

The overall objectives of the MRes course are to:

1. Provide students with a broad understanding the range of urgent environmental challenges facing global society and practical experience of applying AI-based tools to address these challenges. The training programme will be individually tailored to take into account the educational background and interests of the students;

2. Build a cohort of students and equip them with skills that prepare them optimally for PhD research. Students will undertake both individual masters-level research projects, as well as a guided team challenge, before embarking on their PhD research. The aim is to encourage both originality and intellectual independence in tackling complex problems, and to foster team working and leadership skills suited to academic and industrial R&D environments and to policymaking.

3. Develop entrepreneurial and project-management skills and generate awareness of industrial, commercial and policy drivers through relevant cohort activities and close integration of CDT partners in the delivery of the educational programme.

Learning outcomes

By the end of the programme, students will have:

- learnt additional skills in disciplines outside of their first degree;
• gained understanding and command of methods and techniques relevant for research at the interface between artificial intelligence and machine learning on the one hand and the study of environmental change and risk on the other;
• attended lectures in degree level topics bespoke to complement their own strengths and knowledge base upon entry, gaining a broad overview and specific knowledge of environmental data science, shared across the whole cohort;
• developed skills in research methods through the execution of a masters level independent research project;
• developed a full interdisciplinary PhD proposal they can defend in an oral examination and, if successful, embark on from their 2nd year at the CDT;
• gained understanding of the Enterprise landscape relating to environmental data science;
• developed a good transferrable skills base, including science communication skills, as well as a sound grasp of safety and ethics in research;
• learnt to work effectively in teams as well as individually.

Knowledge and understanding

Students will receive a foundation training in the following topics:
• Foundations of Data Science
• Probabilistic Machine Learning
• Environmental Risk

Environmental Data Analysis

Combined with specialist training in two options, at least one of which is from the following:
• Advanced Machine Learning
• Advanced Topics in Machine Learning and natural language processing
• Cloud Computing
• Deep Neural Networks
• Inference
• Machine Learning and Bayesian Inference
• Machine learning and the physical world
• Mobile Robot Systems
• Statistical Learning in Practice

Statistics in Medicine

with the following options also available:
• Atmospheric chemistry & global change
• The Earth System and Climate Change
• Computational Geosciences
• Fluid Dynamics of Climate
• Frontiers of ice core science
• Natural Hazards

Responses to Global Change

Skills and other attributes

The MRes in Environmental Data Science will develop the essential cross-disciplinary collaborations, foster new research activities and maximise impact, by providing a set of transferable skills and wide know-how across the field of environmental risks, setting the research in a framework to consider the societal and ethical dimensions. This is not merely mathematics, engineering, and computer science serving the environmental science community, it is also about having real-world problems to drive further development of novel AI, statistical, and computational methodologies.

It will:
• Build a cohort of students and equip them with skills that prepare them optimally for
PhD research. Students will undertake both individual masters-level research projects, as well as a guided team challenge, before embarking on their PhD research. The aim is to encourage both originality and intellectual independence in tackling complex problems, and to foster team working and leadership skills suited to academic and industrial R&D environments and to policymaking.

- Develop entrepreneurial and project-management skills and generate awareness of industrial, commercial and policy drivers through relevant cohort activities and close integration of CDT partners in the delivery of the educational programme.
- Introduce a range of professional development topics that build the research and personal capabilities of the participants to prepare them for their future paths, as set out in the programme structure.

Programme structure

The MRes course will comprise four components: 1) Introduction, 2) Foundation, 3) Specialisation and 4) Professional Development. The MRes course structure is outlined below. Each student will be paired with one of the two MRes directors as their MRes principal supervisor. Their individual meetings will provide continuous advice and support in all aspects of the MRes including course and project selection, teaching support and general welfare. This support builds on the additional support students will receive from College Tutors and provides a more bespoke sounding board for them to discuss their progress through the course.

<table>
<thead>
<tr>
<th>Programme structure</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introductory courses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foundation courses (Foundations of Data Science, Probabilistic Machine Learning, Environmental Risk)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foundation courses (Env. Data Analysis)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guided Team Challenge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specialisation courses (incl. at least one ‘application domain’)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research project</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional Development workshops</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PhD project preparation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1) **Introduction:** A compulsory intensive introductory week will act as an induction to the MRes and associated CDT, to encourage collaboration across the cohort and to provide a basic introduction to key concepts that will be required for the Foundation courses, accounting for the diverse backgrounds of the students. The five days will comprise three hours of lectures (am), a practical session and final lecture (pm), and evening social activities.

Four introductory courses (5 hours each across the week) will introduce mathematical background:

- **Linear algebra** [including eigenvectors/values, index notation, manipulation of matrices and vectors]
- **Multivariate calculus** [multivariate integration, differentiation, constrained optimisation and Lagrange multipliers, functional derivatives]
- **Probability** [basic distributions and their properties (esp. Gaussians), transformation of variables, mean/variance/covariance, basic methods of parameter estimation]
- A basic introduction to machine learning and data analytics [guiding principles, examples of problems that have been successfully addressed using these approaches]

Each afternoon a three-hour hands-on interactive session will be devoted to a different theme:

- Day 1: Team-based challenges (fun, ice-breaker engineering problem solving)
Day 2: Research skills (incl. access to datasets, availability and use of computing facilities, and introduction to responsible research and innovation)
Day 3: Computer programming methodologies for large-scale computing, related issues & trends
Day 4: Introduction to environmental data incl. GIS and Earth observation datasets
Day 5: Introduction to standard machine learning approaches and tools (e.g. Tensorflow)

2) **Foundation**: The compulsory Foundation course will run through Michaelmas Term (MT) and Lent Term (LT) of the academic year. It consists of a set of mandatory 16 hour lecture courses, supported by examples/tutorial classes and practicals, and a guided team challenge. The compulsory courses comprise: *Foundations of Data Science (MT)*, *Probabilistic Machine Learning (MT)*, *Environmental Risk (MT)*, *Environmental Data Analysis* (LT). The courses will be supported by *examples classes* designed to promote discussion, and where necessary clarification, of concepts introduced in the lectures and the opportunity to reinforce learning by working through example problems. These classes will also be used to revisit the topics presented in the introductory courses. They will include problem-solving sessions, computer-based practicals and in-depth tutorials and will be conducted by experienced teaching assistants. The lecturers and teaching assistants will be available for consultation as required.

The other component of the Foundation course is the **Guided Team Challenge**. Developed in collaboration with the European Space Agency (ESA)’s new innovation lab in Earth Observation (“Ø-lab”), the students will compete in two teams to construct a data-driven evidence base in support of an environmental challenge. A different challenge will be posed for each annual cohort. The project will be on-going through the MRes year, with guided activities to enable the students to access and process Earth Observation data and to use AI approaches to generate decision-support information. ESA will provide all the necessary data, processing tools and computing infrastructure, and access to Ø-lab expertise (see LoS). The teams will write a report and present their results to a judging panel with representatives from the CDT (incl. partners) and ESA. This will build cohesion within the cohort, encourage innovative thinking, provide practical experience that will support later PhD research and encourage collaborative working. It will deliver both technical research skills and a wider range of translational skills.

3) **Specialisation**: Here, the students will choose two lecture courses (each 24 hours) of which at least one must be an “Application Domain” course. These courses will allow students to tailor their training according to their individual research interests and backgrounds. They leverage a range of existing courses (timetabled during Oct-Mar) that are currently dispersed between different departments and bring them together under a common theme. Bespoke support with weekly tutorials provided by CDT teaching assistants will complement the lecture courses and ensure educational material is embedded in a context relevant to the MRes’s focus. New courses will be developed, tailored for the MRes, where there are gaps and sufficient demand, and to meet changing knowledge.

**Application Domain**: at least one from
- Atmospheric chemistry & global change
- The Earth System and Climate Change
- Computational Geosciences
- Fluid Dynamics of Climate
- Frontiers of ice core science
- Natural Hazards
- Responses to Global Change

**Other relevant course include**:
- Advanced Machine Learning
- Advanced Topics in Machine Learning and natural language processing
- Cloud Computing
- Deep Neural Networks
- Inference
- Machine Learning and Bayesian Inference
- Machine learning and the physical world
- Mobile Robot Systems
The other component of the specialisation course will be the **MRes research project**. The MRes supervised research project will provide the students with research experience. It will also further fostering collaboration between disciplines to stimulate new research avenues. During the first term, two afternoons will be dedicated to “shotgun talks” where prospective supervisors will describe their research interests. As well as providing a broad overview of the CDT’s scope, this will allow the students to start to identify research areas of interest. Supervisors will submit proposed MRes research project descriptions for review by the CDT Teaching Committee by December; topics engaging CDT partners will be particularly welcomed. A list of agreed projects will be advertised in January and students will be encouraged to discuss their preferences with potential supervisors. While there will be no formal requirement for dual supervision between the “AI” and “Application domain” sides, it will be expected that supervisors will put in place arrangements to ensure there is sufficient access to the required breadth of expertise. Final approval of research project allocations will be made by the Teaching Committee in late March. Students will start their project in April with a 5000 word scientific report due at the end of July. Students will give a 15 minute oral presentation on their project to the full CDT cohort in August.

4) **Professional Development:** Five compulsory one-day professional development workshops will be held during May-June with extensive end-user involvement in design and delivery. They will include keynote lectures, group activities and discussions. All resources will be made available on the MRes/CDT website. Responsible Research and Innovation (RRI) concepts and considerations will be embedded in each workshop, with three more one-day workshops on key aspects of RRI delivered by RRI-accredited academics.

- **Sustainable Business Workshop:** providing an overview of the current thinking on sustainability in a business context with case studies provided by a number of industry contributors (facilitated by Cambridge Institute for Sustainability Leadership, CISL).
- **Policy Workshop:** bringing together students and civil servants to demonstrate, through worked examples, the value of building links between evidence, expertise and policy making, to create networking opportunities and to share and inspire good practice for evidence-informed policy making (facilitated by Centre for Science and Policy, CSaP).
- **Societal, Legal and Ethical Dimensions Workshop:** exploring the wider implications of research involving the application of AI, developed in collaboration with Cambridge Big Data and the Centre for the Study of Existential Risk (CSER). The workshop will include an introduction to the framework for Responsible Research and Innovation that the students will be expected to follow during the development and implementation of their PhD research projects and associated activities.
- **Entrepreneurship Workshop:** an interactive session in collaboration with Cambridge Big Data and the Centre for Entrepreneurship Learning, hosted at the British Antarctic Survey (BAS) Aurora Innovation Centre, with entrepreneurship training and a mini hackathon-style competition.
- **Impact, Engagement and Communications Workshop:** organised by the Communications Team at the BAS, including guidance and individual feedback on writing effective press releases and briefing notes, tips for presenting to a general audience, and basic media and social-media training.
- **Responsible Research and Innovation Workshops:** delivered with input from the University’s Research Governance and Integrity Officer, three one-day workshops will cover 1) Introduction to Responsible Innovation (incl. mapping student’s proposed PhD research to
AREA framework), 2) Research integrity (incl. introducing the University’s research ethics system), 3) Hands-on RRI workshop (students plan appropriate strategies to address RRI issues).

Teaching methods

Teaching and Learning will be delivered and achieved through a combination of methods. Lectures, examples classes, practical classes and tutorial workshops will form the key delivery of much of the Foundation and Specialisation components. Group work will form a major part of the team challenge, introductory, and professional development components. Individual guided research development will be the main component of the individual research project.

We aim to make the core MRes lectures fully open-access. Each lecture will be live-streamed and video-archived for future viewing on the MRes/CDT website alongside lecture notes. In addition to the benefits of open-access, this will enable students to revisit the lectures as desired. Material from associated examples classes and practicals, including relevant datasets, will be archived alongside the lectures to form a comprehensive on-line learning platform. In a new and innovative departure from traditional methods we will collaborate with Cambridge University Press to deliver teaching materials in an interactive, evolving browser-based format, such as Jupyter notebooks, which also semantically links audio/video of lectures to written materials.

Assessment methods

The examination for the MRes degree shall consist of:

- a report on the MRes research project of not more than 5,000 words in total, exclusive of tables, footnotes, bibliography, and appendices;
- an oral presentation to the cohort on their MRes research project;
- course-work which may include practical/example classes, written work, group work, and class participation;
- written PhD project proposal of not more than two pages;
- an oral examination on the work submitted by the candidate, including the PhD project proposal, and on the general field of knowledge within which such work falls.

The portfolio and oral shall provide evidence to satisfy the Examiners that a candidate can design and carry out an original investigation, assess and interpret the results obtained, and place the work in the wider perspective of the subject. No unseen examinations will be employed.

In order to pass the MRes degree a student must:
Obtain at least ‘pass’ (as defined in the course published description) in each of the following four elements:
1. Coursework
2. Research project
3. Research project presentation
4. PhD project proposal

and pass the

5. Oral examination

A ‘marginal fail’ (as defined in the course’s published description) in any of elements 1-4 above may be compensated by a ‘Distinction’ in another element from 1-4 and a ‘pass’ in oral. Fail on the oral examination may result in a re-viva at the discretion of the examination
panel and CDT Management.

A candidate who does not pass the MRes may have met the requirements for the CPGS and, if so, may be recommended for the CPGS as an exit award.

The results for this MRes degree are pass threshold = 60%, with a Distinction awarded for marks above 75%. The individual marks on the course components are combined as

1. Coursework 10%
2. Research project 65%
3. Research project presentation 10%
4. PhD project proposal 15%

Marginal fail threshold = 55%.

Entry requirements

Candidates should have a minimum of a UK 2.i Honours degree or international equivalent in a relevant degree subject. Shortlisting and interview of applicants will take place in LT of the academic year preceding entry.

Progress requirements

Progression to the (probationary) PhD requires a candidate to pass the MRes examination and fulfil all the following requirements to the satisfaction of the CDT Management Committee:

- received ‘satisfactory’ supervision reports in all three terms;
- satisfactory attendance at compulsory training;
- have the agreement of two participating PIs as the PhD supervisors;
- produced a satisfactory research proposal that lies within the field of enquiry offered by the CDT course which may be within the Dept of Earth Sciences or other partner Department within the University.

Student support

The CDT conforms to the Board of Graduate Studies’ Code of Practice..

Further information about student support may be obtained on request to the Department of Earth Sciences.

Management of teaching quality and standards

High standards of teaching and learning will be maintained by:

- The completion of Annual Quality Updates by the MRes management board, to enable central overview of provision and assist in dissemination of good practice
- Scrutiny of the reports of External Examiners for the MRes
- Encouraging student engagement at both the local level, through involvement in the CDT management committees, and at a central level by participation in the National Student Survey (NSS) and the Student Barometer
- Holding reflective, centrally-coordinated, Learning and Teaching Reviews for all teaching institutions every six years to explore provision and suggest constructive courses of action
- Mentoring, appraisal, and peer review of staff, and encouraging staff participation in personal development programmes
Graduate employability and career destinations

As this MRes forms a part of 1+3 PhD (CDT) programme the MRes graduates are expected to continue towards a PhD degree at Cambridge.

The adsorptive capacity for graduates in this sector is immense: the AI4ER CDT will provide graduates to a sector desperate to recruit far greater numbers of qualified personnel, evidenced by our strong backing from partners.

The Royal Society’s influential report on Machine Learning emphasised it could be a “key enabler for a range of scientific fields” and highlighted current work being undertaken in Cambridge University and British Antarctic Survey using data science approaches to understand the effects of climate change on cities and regions. It concluded, however, that to put machine learning to use, a pipeline of informed users or practitioners is needed. The AI4ER CDT addresses this skills deficit within the societally-important domain of environmental risk. The need for greater training that bridges AI and environmental science was further validated at a Centre for Science and Policy (CSaP) workshop (June 2017) where experts from academia, policy and industry discussed the opportunities for advances in machine learning to be applied to environmental risk assessments. Outside of academia, strong commercial interest is growing in this area, further confirming the need for related skills training. For example, in the past twelve months Microsoft announced a large investment in AI for Earth, two US-based start-ups harnessing AI for geospatial analytics attracted substantial investments.

The Careers Service maintains link with employers and takes their needs and opinions into account in the services which it provides for students. The Careers Service also allocates a Careers Adviser to each College, faculty and department to act as a point of contact.

Every effort has been made to ensure the accuracy of the information in this programme specification. At the time of publication, the programme specification has been approved by the relevant Faculty Board (or equivalent). Programme specifications are reviewed annually, however, during the course of the academical year, any approved changes to the programme will be communicated to enrolled students through email notification or publication in the Reporter. The relevant faculty or department will endeavour to update the programme specification accordingly, and prior to the start of the next academical year.

Further information about specifications and an archive of programme specifications for all awards of the University is available online at: www.camdata.admin.cam.ac.uk