

## **MODS: Mapping Knowledge with Data Science (MSc+PhD/PhD)**

It is exactly 100 years since the enrolment of the UK's first doctoral student, making it an opportune moment to take stock of the economic impact of the PhD. Working in collaboration with the British Library, and fully funded by the ESRC (see [eligibility](#)), you will have full access to two unique data sets covering more than 450,000 PhDs in order to map the flow of ideas within and between institutions using cutting-edge text-mining and network analysis techniques.

At a time of great uncertainty for UK Higher Education, understanding the geography of innovation embodied in doctoral research offers a rare opportunity for a student to tackle a question with fundamental, built-in policy implications: how does ground-breaking work (whether theoretical or empirical) in the natural and social sciences, as well as the humanities, emerge from within particular institutional contexts and how is it disseminated over time and space to other institutions?

### **Research Description:**

Selecting between one and three disciplines/topics for close study (*e.g.* AI, a non-STEM domain such as cultural geography), you will examine how groups and departments are impacted by the departure or arrival of researchers and their research interests. This could form a platform for subsequent post-doctoral work on other domains and policy recommendations leading towards a wider view of 'impact' from doctoral research.

A key challenge, however, is that there are many missing values in the data. We face a data linkage/text mining problem that will require you to think creatively about the pros and cons of different techniques: *e.g.* social network analysis of co-authorship links could help to fill in missing supervisor data, but so could text-mining of thesis texts to find acknowledgements, or scraping of institutional repositories for more recent theses.

The project therefore offers you the opportunity to develop and apply a range of techniques in a comparative context: you will need to evaluate techniques for accuracy and efficiency, and even consider consensus or meta-learning approaches (*i.e.*, employing ML to select the most appropriate technique for individual problems). These skills are highly sought after not only in an academic context, but are also in high demand by industry. A successful project implementation could see your code adopted by the British Library!

### **Person Description & Application**

We need a practically-minded but creative student interested in *both* the social science *and* in the application of data science techniques. You must be [eligible](#) for ESRC funding and will be asked to pursue either the 1+3 (MSc+PhD) or +3 (PhD only) track. *Note:* you are eligible for the 1+3 track **even if you already have a first Masters degree in an unrelated discipline.**

If you are pursuing the 1+3 option, then you would normally be expected to [apply](#) for King's highly-competitive [MSc Data Science](#) prior to beginning your doctoral research. **The cost of the MSc is included in the ESRC award.** Depending on your background, in order to comply with ESRC funding rules we may require you to take the [Theorising Big Data](#) module (or an agreed equivalent) as part of your MSc and to complete specific *short-courses* in the social sciences during the first year of your PhD.

Whatever your background, you will be able to provide specific evidence of your existing technical competency level(s) as part of the selection process; however, we are *also* looking for a student who is passionate about the questions that this research seeks to answer: ***this is not a purely technical 'data science' exercise***, but the application of those techniques to questions with wider social and political relevance!

Successful applicants will most likely have followed one of the three following pathways:

#### **Pathway 1: STEM (*e.g.* Computer Science/Statistics/Bioinformatics)**

You will have an undergraduate degree (or first MSc) in a STEM discipline that involved substantial programming and statistical analysis, but are interested in applying your knowledge to pressing social science questions and to messy, real-world data.

### **Pathway 2: Quantitative Social Science (e.g. Economics, Quantitative Geography)**

You will have an undergraduate degree (or first MA/MSc) in a social science discipline that involved substantive programming and statistical analysis, but require training in more advanced data science techniques. Alternatively, you might have learned to code 'on the side' and used scripts as part of your undergraduate or Masters dissertation but, again, require formal training in data science techniques.

### **Pathway 3: Mix of Academic & Practical Experience**

We will happily consider applicants with relevant professional experience equivalent to a Masters-level degree. In particular, if your background includes both the necessary computer science/programming background *and* at least some postgraduate training in the social sciences (or relevant work experience), it may be possible to apply directly to the PhD (+3 *only*) and to undertake specific 'top up' training in the first year of the PhD;

### **Application Deadlines & Requirements**

Applications for admission to the 2018/19 MSc Data Science programme close on 31 March 2018 and the minimum standard for this programme is a 2:1 or better in a quantitative UG subject and >6 IELTS. Consequently, we aim to select a suitable doctoral candidate by the end of February so that you can be admitted to this Masters programme in a timely fashion, and we have set an **application deadline of 23:59 on Monday February 12 2018.**

Please pay careful attention to the **eligibility requirements** of the ESRC as we are unable to make exceptions to these rules. We are not able to offer exceptions to the residential requirements.

Your application will include:

1. A CV of *no more than 2 pages* highlighting relevant study and work experience.
2. A cover letter of *no more than 2 pages* that addresses your interest in, and suitability for, this research project (see *Person Description* above) as well as a basic indication of how you think you might shape/operationalise the research so as to make it your own. In short: convince us that you are the right person for the job!
3. 2 references, at least one of which will be academic, speaking to your suitability for this research project (see *Person Description* above).

We expect to conduct in-person interviews in late February, so please include in the cover letter any constraints (*e.g.* availability, travel time and/or cost) that might impact our ability to conduct an in-person interview.

Please email Dr. Jonathan Reades ([jonathan.reades@kcl.ac.uk](mailto:jonathan.reades@kcl.ac.uk)) with any questions.