



Company Name:	The Data City
Team / Department:	Data Science
Address:	3rd Floor, Munro House, Duke Street, Leeds LS9 8AG

Provisional title for project:

Developing a database for identifying the most influential directors in the UK using Companies House

Short description of the problem that would be addressed by the project:

The Data City is using A.I to discover the UK's new or emerging sectors and building company lists with accuracy. Companies House will often identify officers (directors) that are registered to multiple companies so directors can be represented using graph theory. It should be noted that, Companies House will often not accurately identify officers that are registered for multiple companies. This calls for a methodology which is able to use additional sources of information to identify when the same officer is registered to multiple companies. The Data City has identified a need for a new unique identifier for officers. Here is an example: *Alex Craven used to be an officer for Bloom Media (UK) Limited but he is now an officer for Data City Innovations Ltd. Companies House does not identify that this is the same person*

Short description of the data sources that would be used in the project, and how they would be used.

The Data City has identified three different sources of data which should help the student. The first is a data set requested by email to Companies House, the second is downloaded from the Companies House website and the third is data that The Data City holds on company directors. We suggest combining these for the most accurate methodology to create the database and display the data

Would any work by the student need to be carried out on site at the Company with the exception of supervisory meetings?

The Data City is based in Leeds and we encourage the student to visit us (COVID-19 guidance permitting) but this is not essential. There is plenty of space in the office if the student chooses to work full time however we cannot guarantee a member of The Data City will be available to supervise the student

Any issues of data confidentiality and IPR that would need to be resolved

Yes

Essential skills

This will involve researching, identifying our customers' needs, creativity with data and python or C#

Desirable skills

Financial/business analysis

Preferred degree programmes (if any)

A demonstrable specialism and background in data and data science

Preferred selection method

Online or telephone interview

Support and training offered by the company

A full range of support and training is available from The Data City, including a 1 hour meeting every 2 weeks and full time access to our data scientist

Financial assistance offered by the company

The organisation will pay the honorarium (£1000). Travel or other expenses reimbursed as appropriate

Any other comments

The Data City will offer supervision from a data scientist who is a former student of the Masters Dissertation Scheme. We encourage potential students to ask questions and make the most out of applying their data interests to a real-world problem. There is a likely to be an opportunity for additional paid work at The Data City throughout the duration of the research. The student will be paid the [national living wage](#) and we expect this will consist of 10-20 hours per week. We will record this with the student using a timesheet to ensure the student does not exceed 20 hours and affect their dissertation. This is optional and this will not have any effect on the candidate's assessment. The student will be most likely supporting The Data City in building lists which represent the emerging economies that do not fit into traditional SIC codes and helping our data analysts develop taxonomies