# **THOMSON REUTERS DATA SCIENCE RESEARCH FUND**

# Big, Open and Linked Data

# GUIDANCE FOR APPLICANTS

## Background

Thomson Reuters is a global organisation providing data and technology services to enable professionals in the financial and risk, legal, tax and accounting, and media markets to make the decisions that matter most, all powered by the world's most trusted news organisation. The company has a long history of connecting, linking and analysing information across a wide variety of sources in a number of different domains.

In 2016, Thomson Reuters established a research fund at the Data Science Institute with a view to engaging in scientific research and development. Currently, research projects are underway in the areas of legal and blockchain.

The TR/DSI management committee are now soliciting applications from across the College for research projects within the field of Big, Open and Linked Data. Knowledge representation remains an important challenge where datasets are typically large, unstructured, or semi-structured lacking clear semantics, are volatile, and highly connected and hierarchical. Networks or graphs provide a natural way to describe relationships between these data and provide an approach for integration from heterogeneous sources. Graph databases are useful for data management and analysis and are particularly efficient for traversal type queries that are important for hypothesis generation surfacing relations between entities that may not have been previously linked. A major requirement of knowledge repositories is to capture contextual information and represent formal statements as subject-predicate-object triples. Resultant knowledge graphs can be explored visually and can be mined under different contexts.

Thomson Reuters have developed a set of tools to enable exploration and analysis of big, open, linked data. [PermID](https://permid.org/) is a unique open identifier for companies, people and financial instruments that facilitate linking data across sources. The Thomson Reuters Intelligent Tagging (TRIT) tool can be used to identify a range of entities in unstructured text and tag these with PermIDs. By using both these tools with a wealth of news articles and curated, authoritative data sets, Thomson Reuters have created a Knowledge Graph of in excess of 5 billion triples, connecting organisations, people, supply chains, financial instruments and quotes. More information can be found here: http://thomsonreutersfinancial.lookbookhq.com/trknowledgegraph/factsheet

We would like to welcome applications that utilise and/or enrich this data.

## Project criteria

The scheme will support research projects:

1. Between 4-12 months
2. In the research area of **Big, Open and Linked Data** using the identified data set
3. With a Principle Investigator must be based at Imperial College and have a permanent contract
4. Up to £100k per project at 100% FEC
5. Staff and non-staff costs

## Funding

A budget of approximately £200k is available for projects running until February 2019. Funds will be awarded at 100% FEC and can only be spent on activities that fit the purpose of this call. Funding can be applied for both staff and non-staff costs. The scheme will not support:

1. Travel costs in excess of £1k
2. New equipment, including staff laptops or servers
3. Any additional administration costs

The funding must be used to support activity at Imperial College only and must remain under the financial control of Imperial College London. All funds must be spent by 30 April 2019.

## IP

TR and IC shall jointly own the Arising Intellectual Property developed under this scheme in accordance with their inventive contribution. Each awarded project will need to complete a Statement of Work, a template of which is included at the end of this document.

## How to apply

Applications should be emailed to [a.ashley-smith@imperial.ac.uk](mailto:a.ashley-smith@imperial.ac.uk)

Applications should have a Principal Applicant based at Imperial and should contain:

* A completed Thomson Reuters Data Science Research Fund Application Form which includes:
  + A written case for support
  + A justification for resources;
  + Relevant experience of applicant and co-applicants
* A one page financial summary: applications should provide information on the requested costs using an InfoEd statement approved by the Department. As InfoEds will not be submitted and will remain as draft, an email confirming departmental approval should also be included in the application, however, proposals do not need to be reviewed by Research Services prior to submission
* CV of the Principal Investigator and Co-Investigators
* For applications that include employment costs, a CV must be provided for any names directly incurred staff.

Deadlines  
Deadline for applications is 12 noon (GMT) on Friday 9th February. Decisions are expected by Friday 23rd February. Projects will be able to commence from 1st April 2018.

## Review and Selection of Applications

All eligible applications will be considered for funding. Under the direction of the Director of the Data Science Institute and the TR/DSI Research Fund Management Committee, proposals will be ranked based on:

* Fit to project criteria above
* Novelty of proposed research
* Relevance to Thomson Reuters interests and strategic priorities

Applications will be informed of the results of the funding meeting via email as soon as possible after the meeting. A formal letter confirming the outcomes will then be sent to Principal Applicants directly by the Chair. Feedback on unsuccessful applications will be limited unless the group decides that additional feedback is necessary or useful.

## Accessing Funds

Information on how to access funds will be provided to successful applicants with their award letter.

## Reporting

Principal Applicants who are successful in receiving an award are required to provide a verbal report on a fortnightly basis.

Each project will be required to complete a template report form which must be completed and returned to the TR/DSI Management Committee at the mid-point of the project and an updated copy on completion of the project. The report should provide details of the funding provided, the results of the research, and how these have addressed the objectives of the call.

Thomson Reuters support through the Research Fund must be acknowledged in any outputs from this funding.

## Contacts

Queries related to research field: [d.mcilwraith@imperial.ac.uk](mailto:d.mciilwraith@imperial.ac.uk)

Queries regarding to funding and application submission: [a.ashley-smith@imperial.ac.uk](mailto:a.ashley-smith@imperial.ac.uk)

STATEMENT OF W**ORK**

|  |  |  |
| --- | --- | --- |
|  | **ITEM** | **DESCRIPTION** |
| **1.** | **TR Reference:** |  |
| **2.** | **Imperial reference:** |  |
| **3.** | **Project/ activity:**  *Choose one of the following:*   * Research Project * Seed Project * Start-up Engagement * Student Enterprise Scheme * Workshops | Research Project |
| **4.** | **Project title:** |  |
| **5.** | **Research Theme to which the Project relates (if applicable):** | Knowledge Graphs |
| **6.** | **Project Start Date:** |  |
| **7.** | **Project End Date:** |  |
| **8.** | **Project Period:** |  |
| **9.** | **Academic Supervisor:** |  |
| **10.** | **TR Supervisor:** |  |
| **11.** | **Student name and department** (if  **applicable):** |  |
| **12.** | **Other Personnel involved in the Project:** |  |
| **13.** | **Project Location and any other specific facilities required for the Project:** |  |
| **14.** | Background Intellectual Property made available by TR (include details of any known third party rights): |  |
| **15.** | Background Intellectual Property made available by Imperial (include details of any known third party rights): |  |
| 16. | Agreed Arising Intellectual Property Option to apply to this Statement of Work: | Option C |
| 17. | TR Field of Use (where Option B is nominated pursuant to section **15** above): |  |
| **18.** | Materials to be provided by a Party for the purpose of the Project: |  |
| 19. | Project Cost (drawn down from the total funding | USD $ |
| 20. | Project Deliverables (including any reporting requirements): |  |
| 21. | Approved by the MB | [YES/NO] |
| 22. | Agreed deviations from the terms of |  |