Last updated: April 2022

## JOB DESCRIPTION AND PERSON SPECIFICATION

| | | | |
|---|---|---|---|
| Post title: | **Research Fellow (Postdoctoral)** | | |
| Standard Occupation Code: (UKVI SOC CODE) | 2119 - Natural and social science professionals | | |
| School/Department: | At Southampton: School of Electronics and Computer Science<br>At Birkbeck: Department of Computer Science and Information Systems | | |
| Faculty: | At Southampton: Faculty of Engineering and Physical Sciences<br>At Birkbeck: School of Business, Economics and Informatics | | |
| Career Pathway: | At Southampton: Education, Research and Enterprise (ERE)<br>At Birkbeck: Postdoctoral Research Assistant | Level: | At Southampton: Level 4<br>At Birkbeck: Research Level 2 |
| *ERE category: | Research pathway | | |
| Posts responsible to: | Dr Thanassis Tiropanis, Professors Adriane Chapman, Alexandra Poulovassilis, George Roussos | | |
| Posts responsible for: | N/A | | |
| Post base: | Office-based | | |

| Job purpose |
|---|
| *Note: a description of the institutions (University of Southampton and Birkbeck, University of London) is annexed at the end of this document).*<br><br>To undertake research in accordance with the ESPRESSO research project under the supervision of the award holders. To undertake leadership, management and engagement activities.<br>Searching over data using Information Retrieval (IR) and Database Querying techniques is key for the development of data-driven services. Techniques that rely on centralised datastores have been thoroughly researched in the past decades. Techniques relying on distributed but centrally controlled datastores have also been researched and deployed. However, there is an emergent paradigm of decentralised services, that relies on the use of distributed personal online datastores (pods). Decentralised services can preserve the privacy of individuals by allowing them to impose access and caching restrictions on their personal data, stored in individual pods that they own and control. Data search in this context presents significant performance and scalability challenges since access rights to datasets and content can vary for different search parties, and different data governance constraints on replicating, storing and processing data can be in place.<br>The EPSRC-funded ESPRESSO project ("Efficient Search over Personal Repositories - Secure and Sovereign") will research, develop and evaluate decentralised algorithms, meta-information data structures and indexing techniques to enable large-scale data search across pods hosted on distributed pod servers, addressing both keyword-based search and SPARQL querying utilising SPARQL endpoints. Our research will be implementation-independent, but will take advantage of existing implemented frameworks and ecosystems (SOLID, NSS, Dataswift/HAT, gtk-gnutella, Gaian DB) for the empirical studies. We will investigate the performance of our new techniques both in generic settings and in the Health domain, which provides a wealth of scenarios presenting individually-controlled data access and caching requirements. The scenario |

development and evaluation will be undertaken in collaboration with industry and academic partners who are engaged in research and development of decentralised services.

Specifically, the research objectives of the EPSRESSO project are:

O1: Design of decentralised algorithms supporting large-scale keyword-based search and distributed querying over datasets and content stored across pods maintained on distributed servers.

O2: Design of access control-aware meta-information data structures and indexes to enable discovery of both structured and unstructured data stored in distributed pods under owner-specified access and caching rights.

O3: Performance evaluation of our algorithms, data structures and indexing techniques, in both generic and health-related application scenarios.

O4: Active engagement with the research community and industry stakeholders.

The project will collaborate with NExT++ centre in Singapore and the SOLID and HAT project teams. We will actively engage with the academic community and industrial stakeholders through dedicated events. The project's findings will inform current research in distributed systems, databases, the digital economy and cybersecurity, as well as the design of innovative decentralised applications, and will help to address the policy challenges relating to data sovereignty and privacy.

We are seeking three Postdoctoral Research Fellows to work on the project full-time for three years: Posts A, B and C. Posts A and C will be based at the School of Electronics and Computer Science, Southampton University. Post B will be based at the Department of Computer Science and Information Systems, Birkbeck, University of London.

Familiarity with the technologies of the SOLID framework will be required for Post A and will be highly desirable for Posts B and C. Familiarity with regulatory frameworks for data access and processing will be required for all three Posts.

All three Postdoctoral Research Fellows will work together on initial platform development and test dataset preparation during the first 6 months of the project. All three Postdoctoral Research Fellows will contribute to the project's engagement and dissemination activities throughout the project (Objective 4 above).

**Post A**

For Post A you will have a background in peer-to-peer systems, distributed query processing and information retrieval. Experience in Linked Data and Semantic Web technologies is also highly desirable.

Post A will be leading the research on decentralised search and querying algorithms over pods (Objective 1 above). This will include algorithms for both 'blind' search (i.e. not having access to the new meta-information data structures) and 'informed' search (having such access), considering both top-k and exhaustive queries. We will also consider caching restrictions on the search algorithms. Also considered will be distributed basic graph pattern (BGP) SPARQL queries,

**Post B**

For Post B you will have a background in distributed information management and information retrieval. Familiarity with semantic technologies, metadata schemas, index design and data summarisation will be expected.

Post B will be leading the research on meta-information data structures and indexes to enable discovery of data stored in distributed pods under owner-specified access and caching rights (Objective 2 above). This will involve design of meta-information data structures to capture dataset profiles and statistics and of indexes over structured and unstructured pod content. The trade-offs involved in maintaining meta-information data structures and indexes on the one hand and improving search performance on the other will be investigated.

**Post C**

For Post C you will have a background in information retrieval, algorithm design, evaluation and optimisation.

Post C will be leading the work on performance evaluation and optimisation of the ESPRESSO algorithms, data structures and indexing techniques (Objective 3 above). This will involve development and evaluation of test scenarios to inform the iterative research of Objectives 1 and 2. Both generic scenarios and application-specific scenarios relating to healthcare will be developed. For the former, pods will be populated with synthetic data and content from Information Retrieval benchmarks, extended with data about owners' access and caching restrictions. For the latter we will employ synthetic data obtained from anonymised datasets of human subjects available on the Web and at the NUS NExT++ centre.

| Key accountabilities/primary responsibilities | | % Time |
|---|---|---|
| 1. | To develop and carry out an area of personal research. | 30 % |
| 2. | Regularly disseminate findings by taking the lead in preparing publication materials for refereed journals and conferences, presenting results at conferences, and exhibiting work at other appropriate events. | 15% |
| 3. | Contribute to the writing of bids for research funding. | |

| Key accountabilities/primary responsibilities | | % Time |
|---|---|---|
| 4. | Investigate models and approaches to test and develop them, individually and in collaboration with colleagues. | 30 % |
| 5. | Collaborate/work on original research tasks with colleagues in other institutions. | 10 % |
| 6. | Carry out administrative tasks associated with specified research funding, for example risk assessment of research activities, organisation of project meetings and documentation. Implementation of procedures required to ensure accurate and timely formal reporting and financial control. if required and occasional undergraduate supervision if required, demonstrating or lecturing duties within own area of expertise, under the direct guidance of a member of departmental academic staff. | 10 % |
| 7. | Supervision of the work of junior research staff. | |
| 8. | Carry out occasional undergraduate supervision, demonstrating or lecturing duties within own area of expertise, under the direct guidance of a member of departmental academic staff. | |
| 9. | Any other duties as allocated by the line manager following consultation with the post holder. | 5 % |

| Internal and external relationships |
|---|
| Direct responsibility to holder of research award or academic supervisor.<br>May have additional reporting and liaison responsibilities to external funding bodies or sponsors.<br>May be asked to serve on a relevant School/Department committee, for example research committee.<br>Collaborators/colleagues within the project, and in other work areas and institutions. |

| Special Requirements |
|---|
| To be available to participate in fieldwork as required by the specified research project.<br>To attend national and international conferences for the purpose of disseminating research results.<br>To understand equal opportunity, ethical and privacy issues and ensure that the research content, data collected and inferred, and methods used are in accordance with these.<br>To carry out tasks that require the learning of certain skills.<br>To engage in continuous professional development.<br>*Applications for Research Fellow positions will be considered from candidates who are working towards or nearing completion of a relevant PhD qualification. The title of Research Fellow will be applied upon successful completion of the PhD in that case. Prior to the qualification being awarded the title of* ***Senior Research Assistant*** *will be given.* |

## PERSON SPECIFICATION

| Criteria | Essential | Desirable | How to be assessed |
|---|---|---|---|
| Qualifications, knowledge and experience | PhD or equivalent professional qualifications and experience in Computer Science.<br>Possess sufficient breadth and depth of specialist knowledge in the discipline to develop research methodologies, conduct research, evaluate research outcomes.<br>Experience with Web development frameworks, the | PhD in distributed database systems, information retrieval, computer networking, or semantic web.<br>Knowledge of standards and approaches related to the redecentralisation of the Web.<br>Experience with Linked Data and Semantic Web technologies (SPARQL RDF, RDFS, OWL, ontology editors). | Application documents/Interview |

| | | | |
|---|---|---|---|
| | Unix programming environment, Java, JavaScript and the C programming language.<br>Familiarity with regulatory frameworks for data access and processing.<br>Hands-on experience of software design, development and performance evaluation.<br>Experience of production of technical documentation, including architectural and design documents, API descriptions, test plans for executing automated test suites and manual tests, software build and installation instructions.<br>Experience of using version and source-control software e.g. GitHub<br>Computer proficiency in standard packages (e.g. word processing, spreadsheets, e-mail and internet use)<br>For Post A: Detailed understanding and knowledge of peer-to-peer systems, distributed query processing and information retrieval. Familiarity with the SOLID (Social Linked Data) framework.<br>For Post B: Detailed understanding and knowledge of distributed information management and information retrieval. Familiarity with semantic technologies, metadata schemas, index design and data summarisation.<br>For Post C: Detailed understanding and knowledge of information retrieval, algorithm design, evaluation and optimisation. | Familiarity with Gaian DB, gtk-gnutella, SOLID-NSS, and Dataswift/HAT frameworks. | |
| Planning and organising | Able to organise own research activities to deadline and quality standards<br>Initiative and creativity to ensure research is effective<br>Excellent communication and inter-personal skills, including the ability to collaborate with others within the team and with the project's range of stakeholders | | Application documents/Interview |
| Problem solving and initiative | Able to develop understanding of complex problems and apply in-depth knowledge to address them<br>Able to develop original techniques/methods | | Application documents/Interview |

| | | | |
|---|---|---|---|
| Management and teamwork | Able to contribute to School/Department management and administrative processes<br>Work effectively in a team, understanding the strengths and weaknesses of others to help teamwork development<br>Commitment to working with diversity | | Application documents/Interview |
| Communicating and influencing | Communicate new and complex information effectively, both verbally and in writing, engaging the interest and enthusiasm of the target audience<br>Able to present research results at group meetings and conferences<br>Able to write up research results for publication in leading peer-viewed journals and conferences<br>Work proactively with colleagues in other work areas/institutions, contributing specialist knowledge to achieve outcomes | | Application documents/Interview |
| Other skills and behaviours | Understanding of relevant Health & Safety issues<br>Proactive in promoting a working environment that is inclusive and engaging; recognising the value diversity brings. | | Application documents/Interview |
| Special requirements | Able to attend national and international conferences to present research results. Able to travel within the UK and abroad several times a year, including spending 2-4 weeks each year in Singapore. | | Application documents/Interview |

**OUR COMMITMENT TO EQUALITY & DIVERSITY**

The University of Southampton is in the top 1% of world universities and in the top 10 of the UK's research-intensive universities. The University of Southampton is committed to sustainability and being a globally responsible university and has recently been awarded the Platinum EcoAward. Our vision is to embed the principles of sustainability into all aspects of our individual and collective work, integrating sustainable development into our business planning, policy-making, and professional activities. This commits all of our staff and students to take responsibility for managing their activities to minimise harm to the environment, whether this through switching off non-essential electrical equipment or using the recycling facilities.

Equality, diversity and inclusion is central to the ethos in the School of Electronics and Computer Science. We particularly encourage women, Black, Asian and minority ethnic, LGBT+ and disabled applicants to apply for these positions. We are committed to improving equality for women in science and have been successful in achieving an Athena SWAN bronze award in April 2020. We give full consideration to applicants that wish to work flexibly including part-time and due consideration will be given to applicants who have taken a career break. The University has a generous maternity policy (subject to qualifying criteria) and onsite childcare facilities.

At Birkbeck, University of London, we are committed to providing the highest quality academic and working environment where all staff, students, visitors and contractors are welcomed respected and treated in a fair, consistent and non-discriminatory manner. The College is proud of its diversity and welcomes applications from all sections of the community. No one will be treated unfairly because of their sex, race, disability, sexual orientation, age, religion or belief, carer status, political belief, pregnancy/maternity, social class, gender identity or marital/civil partnership. The College is a member of the 'positive about disability' Disability Confident scheme and will interview all candidates who both declare a disability within the meaning of the Equality Act 2010 and meet the minimum essential criteria for the post, subject to any limits on the overall number of interviews.

The Department of Computer Science & Information Systems embraces and celebrates the differences between people, recognising the strengths and benefits of a diverse, inclusive society, workforce and student body. In the spirit of our founder George Birkbeck, we believe that education should be available to all sections of society, providing an inclusive working and learning environment for students and staff, so that all may develop to their full potential. We are committed to promoting equality and diversity in all aspects of our activities, through the development of fair and equitable procedures, academic programmes of study, courses, training and development programmes, which are consistently applied and regularly monitored. We encourage and promote an inclusive approach that treats colleagues, students and other service users with respect. We work to ensure that we fully support national schemes, including Athena SWAN, Disability Confident, Mindful Employer, Stonewall Diversity Champions programme and the WISE Campaign. Since April 2018, the Department has held an Athena SWAN Bronze Award in recognition of of our active support for the initative which encourages and recognises the advancement of women within STEM subjects.

## JOB HAZARD ANALYSIS

**Is this an office-based post?**

| | |
|---|---|
| ☒ Yes | If this post is an office-based job with routine office hazards (eg: use of VDU), no further information needs to be supplied. Do not complete the section below. |
| ☐ No | If this post is not office-based or has some hazards other than routine office (eg: more than use of VDU) please complete the analysis below.<br>Hiring managers are asked to complete this section as accurately as possible to ensure the safety of the post-holder. |

## - HR will send a full PEHQ to all applicants for this position. Please note, if full health clearance is required for a role, this will apply to all individuals, including existing members of staff.

## ABOUT THE INSTITUTIONS

**The University of Southampton** (www.southampton.ac.uk) is ranked in the top ten of research-led UK Universities. It has been recognised as an Academic Centre of Excellence in Cyber Security Research, part of a Government strategy protecting and promoting the UK in a digital world since 2012.

**The School of Electronics and Computer Science (ECS)** (www.ecs.soton.ac.uk) at the University of Southampton is internationally regarded for its world-leading and transformative research. ECS has played a leading role in developments in the areas of Web Science, Open and Linked Data, and Data Sharing in Web Observatories. It has driven UK government-level policy decisions that led to the creation of http://data.gov.uk, has launched the http://data.gov.uk initiative with other UK universities, and was instrumental in the creation of the Open Data Institute. The **Web & Internet Science Group** in ECS researches socio-technical systems at scale using data science, computer science and social science methods to understand and engineer the interaction between computational and human constituents of very large scale information systems - the Web and the Internet.

**Birkbeck, University of London** (www.bbk.ac.uk) is a world-class research and teaching institution, a vibrant centre of academic engagement and excellence and London's only specialist provider of evening higher education. There are nearly 19,000 students studying short courses, certificates, diplomas, first degrees, postgraduate taught and postgraduate research degrees. Birkbeck provides Londoners with the unique opportunity to fit study around their busy lives. Founded in 1823 as the London Mechanics' Institute, Birkbeck was incorporated in the University of London in 1920.

Birkbeck's reputation as a world-class research-intensive institution was cemented by the Research Assessment Framework in 2014. Birkbeck was 30th in the UK for research, with 73% of our research rated 'world leading' or 'internationally excellent'. We are home to over 40 research and specialist institutes and among the awards our research has won in recent years are the Queen's Anniversary Prize for excellence in higher education at the Centre for Brain Function and Development. This unique combination of ground-breaking research and innovative teaching provides an inspirational learning experience for our students; we have consistently ranked in the top five in London for teaching and student satisfaction in the National Student Survey.

Birkbeck's main Bloomsbury campus is in the heart of academic London, home to a number of other universities and colleges of the University of London, including University College London and the School of Oriental and African Studies. We are situated among beautiful, leafy squares in one of the world's greatest concentrations of libraries, including the British Library and Senate House Library.

The **Department of Computer Science and Information Systems (CSIS)** (www.dcs.bbk.ac.uk) is one of four departments in Birkbeck's School of Business, Economics and Informatics and one the oldest academic computing departments in the world. It is a world-class centre of expertise in algorithms, data analytics, data management, experimental data science, knowledge representation, and program verification. We engage in a

broad spectrum of both fundamental and applied research, host two research centres, the Birkbeck Institute of Data Analytics and the Birkbeck Knowledge Lab and are a key partner in the national Institute of Coding.

The **Birkbeck Knowledge Lab** (www.klab.bbk.ac.uk) is an interdisciplinary research centre drawing on expertise from the Department, School and College, to investigate how digital technologies and digital information are transforming our learning, working and cultural lives. The Knowledge Lab conducts interdisciplinary research in areas including healthcare, digital economy, digital humanities and technology-enhanced learning, in collaboration with stakeholders from the health, business, cultural heritage and education sectors.