

**Job Posting #: 2-024**

## **POSITION SUMMARY**

### **INDOC RESEARCH – Post-Doctoral Fellow, Data Federation and Query**

---

**Position Title:** Post-Doctoral Fellow, Data Federation and Query

**Company:** Indoc Research

**Posting Date:** November 21, 2017

**Expected Starting Date:** January, 2018

**Location:** Toronto, ON or Kingston, ON

---

#### **Job Summary:**

We are seeking a self-starter with good experience and great passion in data federation and cross-modality data query.

You, our ideal candidate, are looking for a small and dynamic multidisciplinary team of researchers and engineers who work together with minimal hierarchy and red-tape. You are prepared to bring your drive, your experience, and your passion to contribute at all levels to the entire team. We, Indoc Research, are that team. We build and manage complex medical research infrastructure for collaborators and clients. We bring together prominent research organizations across the country, including the Neuroimaging Group at the Rotman Research Institute, the Centre for Advanced Computing at Queen's University, and the Electronic Health Information Laboratory at Children's Hospital of Eastern Ontario. Together we have created large scale informatics platforms involving diverse and complex data modalities (e.g. imaging, proteomic, genomic, clinical assessments) across multiple disease areas (e.g. neurodegeneration, depression, cancer).

In the FoRCE (Focus on Research and Clinical Evaluation) project, you will have the opportunity to stretch and develop your data federation and data query skills to a whole new level. FoRCE will require addressing complex data types in real-time, providing responsive and innovative federation, query and visualization of high-dimensionality data, while maintaining security and privacy of data in a shared clinical and research setting. FoRCE will address critical gaps and fulfill currently unmet and urgent needs in both clinical and research communities, acquiring, federating and sharing data from distributed settings such as critical care units, clinical laboratories and hospital imaging facilities. FoRCE will enable real-time streaming of medical data, support novel cross-modality data fusion,

while providing regulatory and ethics compliance as required in these clinical settings. As the post-doctoral fellow on this project, you will have a critical role at all levels of this project, in design and implementation, architecture and testing, deployment and support. And you will be building the future of medical science.

This Post-doctoral fellowship position will provide the successful candidate with opportunities to acquire hands on experience in technical aspects of development and maintenance of large scale research informatics systems, within industry and academic settings.

#### **Key Responsibilities:**

The Data Federation and Query PDF will be responsible for the following development activities related to the FoRCE project, as well as the wider Indoc Informatics platform as required:

- Development and optimization of data federation subsystem
- Development and implementation of data processing and handling pipelines
- Development of a cross-data modality query engine and associated visualization tools
- Database design (relational and non-relational databases)
- Database administration

#### **Required Qualifications:**

- Ph.D. in Computer Science, Software Engineering or Equivalent

#### ***Data Architecture:***

- Deep understanding in open source based data solutions, including data warehousing, distributed data processing / search engine
- Experience of data federation / integration
- Experience in ETL tools and making data services
- Experience of Big Data ecosystem, including SPARK, Kafka, Hadoop etc.
- Experience of Docker and / or virtualized computing infrastructure

#### ***System Administration***

- Strong experience in Linux server administration
- Familiar with SPARK, Hadoop, Cassandra, Docker Swarm, ELK
- Experience of OVM and docker deployment
- Experience maintaining SGE / PBS clusters

#### ***General Scientific Programming***

- Strong experience in one of the scientific programming languages, Python, R, Scala, etc
- Experience with XML, JSON, parquet
- Good understanding in data processing pipelines and data driven applications
- Experience in versioning control system, Git, SVN. etc.



***Other Skills***

- Strong written and oral communication skills
- Track record of initiative and self-organization
- Willingness and ability to work on multiple projects at the same time
- Solid time management skills
- Demonstrated ability to work within a collaborative team across multiple disciplines
- Willingness to research and learn new approaches and technologies

**How to Apply:**

Please submit your resume and cover letter to **info@indocresearch.org** with the subject title "FoRCE Data Federation and Query PDF Application".

**Suggested deadline:** December 22, 2017. The position will remain open until filled.

indocresearch.org, 258 Adelaide Street East, Suite 200 Toronto, ON, Canada M5A 1N1 W:  
416.703.0311 E: info@indocresearch.org