Dear Prof. Dr. Martin Theobald,

I am writing this statement to express my interest in the postdoc opening that has been announced on DBWorld. I am in my final year of Ph.D. and would be defending in June this year from Universite Libre de Bruxelles (ULB), Belgium & Universitat Politecnica de Catalunya (UPC), Spain. The thesis topic is about temporal graph mining and large graph distributed processing, and it is advised by Prof. Toon Calders from ULB and Prof. Alberto Abello from UPC. As part of my thesis, I have worked on different stream mining data structures, temporal graph analysis, information flow mining in interaction networks, and distributed graph processing. The main focus of the thesis is to design and develop efficient algorithms to analysis temporal graphs and provide distributed processing for such algorithms.

I have been granted the MSc (2008-2011) in Computer science from CMI (Chennai Mathematical Institute), India. ). My masters was part of a joint program conducted by TCS (Tata consultancy services) and CMI wherein I use to go to university for my Master in Computer science course 2 days a week and work as research and training assistant in TCS for rest of the time.

Concerning my industry experience, I have worked both as a researcher in Industry lab working on prototype projects and also as a technical team lead and architect to take one of the prototypes to market as a hugely successful large-scale system to be used by multiple clients from Education domain. I was the lead Java developer and architect for the design and development of the software platform on a multi-tenant architect for Software as a Service model. For the last 4 years, I am working with C++, Python, and Scala to develop open source code to provide the implementation of algorithms and data structures developed as part of our research.

As part of my future research directions I am looking into continuing my work on distributed graph processing for time evolving temporal graphs. Currently, none of the large data processing systems such as Gelly-Flink or Spark-GraphX supports handling of temporal graphs. As part of the research involved in supporting large graph distributed processing I worked on a cost model based approach to determine the best partitioning strategy for a graph on Spark GraphX.

I have been working on graph processing and big data systems as part of my Phd and the post doc position advertised on your group seems a perfect fit for my future research aspirations. I have the necessary programming background in Java and Scala, the Systems background in Hadoop and Spark, and the theoretical background in graph mining and graph processing to contribute to the research project on large-scale extraction, indexing and querying of semantic data in a distributed graph engine.