***Benchmarking and Evaluation of Graph Data Systems***

**Proponente: Mtro. Víctor Ortega, ITESO/UABCC,** [**vortega@iteso.mx**](mailto:vortega@iteso.mx)

1. **Descripción general del proyecto propuesto**

En la actualidad existen estudios que realizaron un análisis comparativo de bases de datos basadas en grafos (BDG), aunque sus resultados son útiles la mayoría no son comparables, ya sea por utilizar distintos *sets* de datos, realizar las pruebas en ambientes distintos, tener preguntas de análisis diferentes o simplemente por tener características no hacen posible homogenizar los resultados obtenidos.

El objetivo del proyecto es realizar un análisis formal de varias BDG en ambientes de prueba similares, con los mismos datos de entrada y preguntas con la intención de ofrecer información acerca de las fortalezas, áreas de oportunidad y el escenario idóneo para utilizar cada una de las herramientas electas.

Las principales etapas del proyecto son:

* 1. Revisar el estado del arte
  2. Seleccionar las BDG a analizar
  3. Generar un ambiente de pruebas para cada BDG
  4. Instalar las BDG
  5. Seleccionar un *set* de datos robusto para las BDG
  6. Diseñar las pruebas
  7. Ejecutar las pruebas
  8. Escribir un artículo con los resultados y el TOG

1. **Vinculación o Colaboración**

Se trabajará de forma indirecta con el Dr. Francisco Cervantes y el Dr. Luis Fernando Gutiérrez

1. **Asignaturas de la MSC relacionadas con el desarrollo del proyecto**
   1. Matemáticas Avanzadas para Computación,
   2. Análisis y Diseño de Algoritmos
   3. Programación Avanzada
   4. Bases de Datos Avanzadas
   5. Sistemas Distribuidos
   6. Manejo y Análisis de Información Masiva
   7. Aplicaciones y Servicios en la Nube
   8. Aprendizaje Automático
2. **Estudiantes participantes en la propuesta**

Dos estudiantes

1. **Bibliografía relacionada**
   1. B. Elser, A. Montresor, An evaluation study of big data frameworks for graph processing, in IEEE Big Data (2013)
   2. Y. Guo et al., How well do graph-processing platforms perform? An empirical performance evaluation and analysis, in Proceedings of Parallel and Distributed Processing Symposium (2014)
   3. M. Han et al., An experimental comparison of prequel-like graph processing systems. PVLDB7(12) (2014)
   4. Y.Lu, J. Cheng, D. Yan, H. Wu, Large-scale distributed graph-computing systems: an experimental evaluation. PVLDB 8(3) (2014)
   5. R. MsColl et al., A performance evaluation of open source graph databases, in Proceedings of PPAAW (2014)
   6. N. Satish et al., Navigating the maze of graph analytics frameworks using massive graph datasets, in Proceedings of SIGMOD (2014)
   7. Y.Zhao et al., Evaluation and analysis of distributed graph-parallel processing frameworks. J Cyber Secure. Mobil. 3(3) (2014)
2. **Víctor H. Ortega Guzmán**

Maestro en Computación (UNIVA) con especialidad en Tecnologías para *Business Intelligence: Data warehouse y Data Mining* (UOC, España) y otra en Mejora de Procesos de Negocio (ITESO). Consultor en el Centro de Gestión de la Innovación y la Tecnología (Ceginnt). Trabaja en proyectos relacionados con las áreas de bases de datos, *business intelligence*, y bases de datos basadas en grafos. Actualmente es estudiante del Doctorado en Ciencias de la Ingeniería en la línea de Software de Alto Desempeño (ITESO).