Estado del arte

*Answering geospatial queries over relational data*

En este *paper* los autores parten del hecho de que los datos geoespaciales son comúnmente almacenados en sistemas manejadores de bases de datos, *DBMS* por sus siglas en inglés, del tipo geoespacial. Para llevar a cabo dicha tarea, se deben de convertir esos datos en datos RDF y almacenarlos en *triple store* cada vez que nuevos datos llegan. Esta labor resulta ser monótona, generando apatía en los administradores de dichas bases de datos para actualizarlas. Si bien existe el paradigma de administración de datos denominada Acceso a Datos Basado en Ontologías, *OBDA* por sus siglas en inglés, que se encargan de ofrecer consultas SPARQL en formato SQL, no existe soporte para datos geoespaciales; esto implica no hacer consultas actualizadas o no obtener resultados de consultas geoespaciales. La solución por parte de los autores de habilitar consultas del tipo GeoSPARQL *en el aire* al no convertir los datos consultados a RDF y después almacenarlos en un *triple store*. Como resultado, se pueden hacer consultas sin tener que hacer conversiones y almacenamiento, solo puras consultas.

*Benchmarking Commercial RDF stores with Publications Office Dataset*

Los autores presentan un *benchmark* para RDF *stores* usando *datasets* de la oficina de publicaciones, *PO* por sus siglas en inglés. En la comparación se miden 4 características: *bulkloading*, escalabilidad, estabilidad y ejecución de consultas. De la misma *PO* se utilizaron *datasets* normalizados y no normalizados. Usando lo mismos *datasets* se construyeron nuevos datos para medir la escalabilidad; en estas consultas de dividieron en 2 categorías: consultas instantáneas, consultas que involucran operadores primitivos, y consultas analíticas con el propósito de medir la calidad de servicio y no solo la velocidad. Se concluye que el rendimiento no es homogéneo entre sistemas y que la calidad y velocidad de resultados dependen de diversos parámetros como el tipo de consultas, características de base de datos o *hardware* utilizado.

*Geographica: A Benchmark for Geospatial RDF Stores*

Se consideró por parte de los autores que no existía un *benchmark* que evaluara *triple store* geoespaciales que fuese usado de una manera conocida por lo que el propósito de haber desarrollado el *benchmark* *Geographica* fue contribuir al estado de arte. En *Geographica* se usaron datos sintéticos y datos del mundo real para probar la funcionalidad ofrecida y el rendimiento de RDF *stores* geoespaciales; se usaron en específico Strabon, Parliament y uSeekM los cuales eran los más completos para la evaluación. El desarrollo de esta comparativa fue con el fin de ofrecer una metodología que permitiera evaluar RDF *stores* de una mejor manera que propuestas previas. Ellos usaron 2 *workloads* de datos: sintético y del mundo real. Con ellos evaluaron eficiencia de funciones primitivas espaciales y el rendimiento de los RDF *stores* en *reverse geocoding, map search y browsing.* Los autores buscan, como trabajo futuro, expandir el *benchmark* para cubrir el estándar GeoSPARQL de una manera completa y probar a *Geographica* en un ecosistema centralizado y distribuido.