Abstrak

Data semistruktural kini mulai diminati oleh sebagian kalangan. Sifat data semistruktural yang fleksibel dalam pendefinisian skema, memudahkan pengembang untuk segera menjawab tuntutan aplikasi dengan cepat. Fleksibilitas skema yang ditawarkan juga sesuai untuk menyelesaikan beberapa permasalahan, terutama untuk domain data yang tidak relasional.

Salah satu model basisdata semi struktural yang populer adalah basisdata berorientasi dokumen. Pada basisdata ini, antara data dan strukturnya dienkapsulasi dalam suatu dokumen. Salah satu alasannya, adalah agar lebih sesuai dalam memodelkan data di dunia nyata, dimana data masih belum dinormalisasi. Konsep dimana data dan struktur dienkapsulasi dalam satu dokumen ini dikenal dengan istilah *self contained data*.

Selain perbedaan model datannya, basisdata dokumen juga memiliki perbedaan dalam melakukan query. Karena basisdata ini masih belum standar, dan masih aktif dikembangkan oleh beragam pihak, maka model querynya pun berbeda antara satu dengan lainnya. Beberapa basisdata dokumen menggunakan bahasa deklaratif, sementara yang lain menggunakan bahasa prosedural. Salah satu model prosedural akan diteliti pada tugas akhir ini, yaitu MapReduce.

Pada tugas akhir ini, akan dilakukan migrasi data *employees*, yang sebelumnya menggunakan basisdata relasional ke basisdata dokumen. Migrasi dilakukan untuk lebih memahami perbedaan antara model relasional serta model dokumen. Setelah itu, akan dilakukan *benchmark* untuk mengukur perbedaan performa antara basisdata relasional dan dokumen.

Hasil dari tugas akhir ini adalah aplikasi yang mampu melakukan migrasi dari tabel relasional ke dokumen dengan tiga model yang berbeda: Self Contained, Normalized, dan Mixed. Selain itu, dari hasil pengujian yang dilakukan diperoleh bahwa fleksibilitas dari data semistrukutral berdampak pada konsumsi *disk space* yang lebih besar, waktu pembuatan indeks yang lebih lama, serta batasan dalam hal query yang bisa dilakukan melalui model MapReduce.

Kata kunci : *Semi Structured Data, Data Migration, MapReduce, Document Oriented Database,Relational Database, Database Benchmark.*