Escuela Colombiana de Ingeniería

Maestría en Informática

SGBD

Laura Milena Ramos

Javier Olarte

**Objetivo:** Obtención del modelo de datos de Cassandra mediante una herramienta tal como lo hicimos con SchemaSpy para Teradata

**Desarrollo:**

Tras validar que SchemaSpy no es la herramienta indicada para obtener el modelo de datos de Cassandra porque el modelo de Cassandra requiere que las consultas sean hechas en cqlsh que es el lenguaje de queries especifico de Cassandra.

De otra parte encontramos la herramienta Hackolade que mediante un proceso de Ingeniería inversa de las bases de datos Cassandra incluye la ejecución de instrucciones cqlsh DESCRIBE para descubrir espacios clave, tablas, columnas y sus tipos e índices.

Si se detecta JSON en columnas de cadena, Hackolade realiza un muestreo estadístico de registros seguido de una inferencia probabilística del esquema del documento JSON.

La conexión de Hackolade a Cassandra se establece mediante una cadena de conexión que incluye la dirección (IP) y el puerto (normalmente 9042), y la autenticación mediante nombre de usuario/contraseña, si procede.

A continuación el keyspace system\_schema desde el cual Hackolade ejecuta la consulta cqlsh para obtener el modelo en Cassandra.



