

Caso de Estudio Base de Datos

“Kickstarter”

Andrea Figueroa R. Javier Jeria M.
abfiguer@alumnos.inf.utfsm.cl javier.jeria.13@sansano.usm.cl

Camilo Rivas F. Camilo Valenzuela C.
camilo.rivas@alumnos.usm.cl camilo.valenzuela@alumnos.usm.cl

Cecilia Reyes C.
reyes@inf.utfsm.cl

7 de mayo de 2016



Introducción

Kickstarter es una plataforma que provee a artistas, músicos, cineastas, diseñadores y otros creadores los recursos y el apoyo necesarios para llevar sus ideas a la realidad. Hasta la fecha, decenas de miles de proyectos creativos de todos los tamaños han podido ser realizados gracias al apoyo de la comunidad de Kickstarter.

Para el siguiente caso de estudio, nos enfocaremos en construir un modelo de datos que sea capaz de resolver el problema del almacenamiento de datos sobre usuarios, proyectos, etc, a partir de las vistas entregadas las cuales deben ser normalizadas, además se requiere un análisis de crecimiento de base de datos y de alternativas de almacenamiento.

Requerimientos

Formar un equipo de 5 o 6 integrantes y realizar las siguientes tareas:

- Registrarse en Kickstarter y explorar la plataforma para lograr familiarizarse con la realidad de la base de datos a diseñar.
- Diseñar un Modelo de Datos Lógico, expresándolo como un conjunto de entidades en 3NF, y a través de un diagrama modelado con MySQL Workbench. En concreto se pide:
 - Normalizar cada vista adjunta a este documento. Por cada vista, justifique qué ítems son necesarios normalizar, y cuáles son ajenos a este proceso. Además, por cada vista justifique si es posible o no el normalizar a 4NF.
 - Identificar atributos derivables.
 - Integrar resultado de la normalización (tenga cuidado con los sinónimos).
 - Proponer posibles mejoras (los atributos derivables no cuentan como mejoras).
 - Graficar Modelo de Datos Lógico (incluidas las mejoras) usando MySQL Workbench.
- Analizar las implicancias de mantener las imágenes y vídeos en la base de datos o mantener los archivos en un sistema de almacenamiento auxiliar. Justifique con un lenguaje técnico su decisión. En el caso de que decida utilizar almacenamiento auxiliar, explique los cambios que se deberían realizar en el modelo de datos (no es necesario graficarlo, puede explicar con palabras).
- Investigue y explique brevemente las tecnologías que actualmente ocupa Kickstarter para almacenar todos los datos, incluyendo datos de la plataforma, objetos multimedia, audios, imágenes y sistemas de caché. Haga un diagrama ad-hoc para representar la situación actual de Kickstarter explicando cómo se relacionan las diversas tecnologías de almacenamiento (en términos sencillos: dibuje la arquitectura de software que usted cree que Kickstarter ocupa actualmente). Mencione el por qué es importante utilizar caché.
- Realizar un estudio de crecimiento de tablas, para ello considere que Kickstarter desde el 2009 a la fecha, ha conseguido 11 millones de usuarios y ha contribuido con 2,3 miles de millones de dolares a la realización de 104.790 proyectos.
Usando estos datos y sus estimaciones de espacio ocupado por registro (fila) en cada tabla, estime cuánto ocupará en GB la base de datos relacional tras 3 años. Esta es una pregunta técnica, por lo que debe ser justificada con un lenguaje y análisis técnico. Suponer cantidades faltantes.
- Conclusiones técnicas y recomendaciones que puedan obtener tras el trabajo realizado. (hint: ¿Basta sólo con usar bases de datos relacionales?)

Vistas

- Vista 1: Acerca del Proyecto. Presentación general sobre el proyecto.
- Vista 2: Actualizaciones.
- Vista 3: Comentarios.
- Vista 4: Categorías.
- Vista 5: Subcategorías.
- Vista 6: Proyectos por subcategorías.
- Vista 7: Nuevo proyecto. Vista compuesta por distintas imágenes (2).
- Vista 8: Registro Usuario.
- Vista 9: Favoritos.
- Vista 10: Recomendados.

Observaciones

- Una vez que tenga el modelo completo y normalizado, debe hacer una navegación por las tablas para comprobar que el modelo obtenido es capaz de obtener los requerimientos (vistas) del caso. No debe presentar este estudio, pero podría ser una pregunta en la presentación.
- El caso deberá ser desarrollado por equipos de **5 o 6 personas**, las cuales deberán hacer un desarrollo acabado de todos los requerimientos exigidos en este documento. Se pondrá especial énfasis en el estudio final, en el cual se pide el máximo de profesionalismo al momento de redactar sus aseveraciones.
- Se premiará orden y presentación.
- El documento debe ser desarrollado en términos técnicos y formales.
- Si utiliza material anexo para el desarrollo, estudio y/o conclusiones finales, debe adjuntar la correspondiente bibliografía: citando a los autores, fuente del medio y fecha de su publicación. En este punto se recomienda cautela respecto de la información que pueda encontrar en la web, es decir, utilice solamente fuentes recomendadas.
- Cualquier copia, será sancionada con nota cero. Tanto si es interna a los equipos, o si se demuestra copia literal de párrafos desde otras publicaciones (copy, paste).
- Se pueden generar supuestos sobre la plataforma, pero estos se deben identificar en el informe.

Informe Escrito

- La presentación del informe es MUY IMPORTANTE. En este ítem debe tener en cuenta, el papel, la letra, el tamaño, la legibilidad del documento, la expresión escrita, hojas numeradas correctamente presentadas, con márgenes claramente definidos, y mucho cuidado en la ortografía.
- El modelo de datos es una parte “fundamental” de su entrega, por lo tanto, una vez finalizado el análisis, verifique que el modelo que proponen, en verdad sea capaz de resolver el problema.
- Todo error de análisis del caso que contenga faltas en lo que respecta a la normalización será considerado grave y por ende sancionado con descuento en nota.
- Evite explayarse demasiado en escritos que no sean estrictamente necesarios, el formato es el de un documento técnico, pero debe ser agradable de comprender.
- Evite espacios o amplias secciones en blanco que denoten la poca eficiencia del uso de las hojas de su proyecto. Evite las hojas en blanco.
- En portada, deben venir todos los datos personales de los integrantes de su equipo. Deben identificarse con un nombre de fantasía para el equipo.

Presentación

- Para la fecha de la presentación usted debe traer su modelo, de manera física en tamaño gigante (mínimo del tamaño de una cartulina), en donde se puedan apreciar las tablas, sus atributos, las asociaciones destacando claves primarias y foráneas. Puede imprimir las tablas y dibujar las asociaciones. Además, debe traer su modelo de datos en formato digital, para poder presentarlo utilizando el proyector (más detalles abajo).
- A cada equipo se le asignará un número, y deberán pegar su modelo tamaño gigante en un lugar asignado de la sala, luego a cada equipo se le asignará un contra-equipo, entre ellos se realizará una evaluación cruzada, indicando los errores que el equipo contrario realizó al plantear el modelo, la idea de esto es evaluar su capacidad crítica a la hora de evaluar un modelo de datos existente. Estas anotaciones se dejarán indicando “tabla y error” a través de un post-it. Para este punto tienen 10 minutos.
- Tras la evaluación cruzada, los equipos volverán a sus puestos originales para desarrollar la defensa ante los errores encontrados. Para este punto tienen 10 minutos.
- Tras la formulación de la defensa, cada equipo escogerá un integrante que actuará como el “elegido” para la exposición. En esta exposición, el “elegido” tendrá 5 minutos para explicar los problemas encontrados y la solución o explicación ante ello. ¡La nota de todo el equipo está en juego! Por lo que algún integrante del equipo puede complementar las respuestas del “elegido” para así mejorar la defensa.
- En cada presentación el staff del curso puede realizar una pregunta extra relacionada con el modelo de datos, esta defensa se debe realizar de manera inmediata, se evaluará la creatividad y el criterio técnico a la hora de responder.
- Finalmente se realizará una conclusión y cierre de este evento.

Evaluación

- El informe escrito vale un 60 % y la presentación un 40 %. La nota final de este caso se multiplicará por la cantidad de integrantes. El equipo deberá posteriormente dividir el puntaje total entre los integrantes del equipo. Ejemplo: Si el proyecto tuvo un 80 y eran 5 integrantes, entonces se tienen 400 puntos. Los integrantes tendrán la decisión final acerca de cómo dividir los puntos, pudiendo dar más nota a quien trabajó más en el proyecto y menos a quien no trabajó lo suficiente.

Consideraciones de Entrega

- La fecha de entrega del informe escrito es el día Jueves 30 Junio, 15.40h horas a través de Moodle.
- La fecha de la presentación del Caso es el día Sábado 2 Julio, 11.30hs.
- Cualquier duda sobre el informe o presentación enviar mail a: abfiguer@alumnos.inf.utfsm.cl.