

Plataforma Web para acceder a diferentes Sistemas de Gestión de Bases de Datos Relacionales

Muñoz Roberto Miguel, Maldonado Calixto, Damiano Luis Esteban, Romero
María Soledad, Cuevas Juan Carlos, Quinteros Sergio Ramón, Guevara
Andrea, Carrasco Agustín

Grupo de Investigación, Desarrollo y Transferencia de Sistemas de Información
Departamento Ingeniería en Sistemas de Información
Facultad Regional Córdoba / Universidad Tecnológica Nacional
Maestro Marcelo López esq. Cruz Roja Argentina – Ciudad Universitaria - Córdoba
0351 - 4686385

{robertmunioz, calixtomaldonado, luis.damiano, romeroma.soledad, juancarloscue, ser.quinteros,
andrezza77, asermax}@gmail.com

Resumen

El propósito del proyecto es estudiar, investigar, diseñar y proponer una Plataforma Web para colaborar en la capacitación y ejercitación del lenguaje SQL, accediendo a bases de datos relacionales.

Es una herramienta de ayuda a la práctica personal, donde el usuario encuentra un entorno con estructuras de datos, tipificación de sentencias y consignas a resolver, para recibir una devolución que le permite crecer en aprendizaje.

Las especificaciones del software de interfaz y del middleware se tomarán a partir de necesidades de usuarios, que inicialmente serán los estudiantes de la cátedra de Gestión de Datos, del tercer nivel de la Carrera de Ingeniería en Sistemas de Información.

El middleware ofrecerá a los usuarios la conexión a distintos motores de base de datos, permitiendo aplicar el lenguaje SQL según los distintos productos que puedan incluirse.

Además brindará un módulo de gestión para usuarios administradores, con privilegio de edición, así pueden

definir estructuras de trabajo, ejecutar scripts con tablas que contengan datos y agregar consignas para ejecutar sentencias SQL.

La arquitectura será web, para que el usuario no necesite instalar software cliente en su equipo y le baste con un browser.

Palabras clave: Base de datos, modelo relacional, SQL, Plataforma web, Software como Servicio, SaaS

Contexto

El proyecto ha sido homologado por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional, con un tiempo de realización desde el 1 de enero de 2016 y por 24 meses. La Unidad Científico-Tecnológica donde se desarrolla es el GIDTSI (Grupo de Investigación, Desarrollo y Transferencia de Sistemas de Información), de la UTN, Facultad Regional Córdoba.

El origen del proyecto se centra en observaciones y encuestas realizadas en la

Cátedra de Gestión de Datos, perteneciente al tercer nivel de la Carrera de Ingeniería en Sistemas de Información, donde un alto porcentaje del cronograma de cursado se destina al estudio y ejercitación del Lenguaje SQL.

El grupo de trabajo está conformado inicialmente por siete docentes investigadores de la carrera, con la colaboración de un graduado y dos estudiantes becarios, de la misma carrera.

Los docentes se desempeñan en cátedras relacionadas a las siguientes temáticas: Programación, Base de Datos y Sistemas de Información.

El equipo de investigación tiene experiencia profesional y en docencia, pero se destaca que mantienen continuidad en proyectos de investigación relacionados a la temática:

- TecnoDB - Administrador de Base de Datos Relacional (2007) [1]
- PROMETEO - Desarrollo de un método y una herramienta para el aprovechamiento de Metadatos de Base de Datos Relacionales (2010) [2]
- Análisis y aplicación de metodologías para la generación de consultas complejas utilizando esquemas OLAP (2010) [3]
- Generador Automático de Modelos de Datos Normalizados en Bases de Datos Relacionales (2015) [4]

Cabe destacar que los integrantes: Ing. Roberto Muñoz, Ing. Calixto Maldonado y el Lic. Luis Damiano, participaron junto a los Ingenieros Enrique Reinoso y Maximiliano Abrutsky, en la escritura del libro "Base de Datos" [5].

Introducción

La temática abordada está relacionada con las bases de datos relacionales, que si bien inicialmente sólo se las utilizaba para persistir datos desde las aplicaciones de las empresas [6], hoy son utilizadas para la toma de decisiones a partir de la

exploración de las bases de datos, con la Minería de Datos [7, 8].

Se considera un tema relevante, dado su impacto en la cantidad de horas hombre destinadas a la resolución de temas inherentes al manejo de datos, en el entorno de bases de datos relacionales. Más aún si se considera que es una tarea componente de la mayor parte de los proyectos de software y que de ella dependen otras como el desarrollo y el testing. Todo desarrollador de aplicaciones debe conocer SQL, para embeber el código necesario en sus programas y así acceder/manipular los datos que solicita un usuario final [9, 10, 11].

Los docentes de la Cátedra de Gestión de Datos observaban que no todos los estudiantes lograban ejercitar en sus equipos personales y sólo lo hacían en el momento de la ejercitación práctica guiada por el auxiliar docente.

Buscando lograr que el entrenamiento desde la cátedra sea eficaz, los docentes fueron proponiendo e incorporando medidas a través del tiempo:

- Actualización permanente de la Guía de Trabajos Prácticos de Gestión de Datos, con ejercicios combinados y revisión permanente.
- Creación de script con la generación del modelo de datos "Empresa" y la carga de datos, para poder ejercitar en Oracle. Se instaló en las aulas del Laboratorio de Ingeniería en Sistemas de Información (LabSis), para que los estudiantes resolvieran ejercicios.
- Con el tiempo se decidió ampliar el alcance del script del modelo, para ser ejecutado en los motores: MySQL, PostgreSQL, DB2 y SQL Server. Estos fueron publicados en el repositorio del LabSis para los estudiantes.
- Posteriormente se creó el script con la estructura y los datos que figuran en la bibliografía de la asignatura, para ser

usado en SQLServer. Ese script también fue publicado.

- Un docente generó una máquina virtual con: sistema operativo, DBMS y los scripts de generación de estructuras y carga de datos.

Esto muestra que el tema de la ejercitación de SQL se analiza de manera permanente, pretendiendo motivar a los estudiantes para una mejor ejercitación y dejarles buen conocimiento de SQL.

A pesar de lo señalado, se continúan detectando inconvenientes en los procesos de aprendizaje de los estudiantes.

La consecuencia de no practicar lo suficiente es que no logran la detección temprana de errores u omisiones, que sólo se alcanza con la resolución de ejercicios.

En el año 2013 la cátedra de Gestión de Datos implementó una encuesta, para conocer las opiniones de los estudiantes y lo relacionado a aspectos de la práctica de SQL. Fue respondida por 219 estudiantes, de los 350 inscriptos. Sólo el 8% de los encuestados indicaron que la práctica era Excelente. Se entiende que los estudiantes también reconocen que la práctica podría mejorarse. Además se verificaron inconvenientes que suelen tener:

- Incompatibilidades entre sistema operativo y DBMS. Es común que los estudiantes consigan el producto a instalar y el proceso de instalación suele no ser exitoso.

- Inconvenientes entre versiones de software de base de datos y script, por ejemplo problemas en formato de fechas.

- Falta de experiencia en la instalación de software. Un estudiante respondió: “Instalé SQL Server pero no pude ingresar por el usuario”.

- Dificultad para conseguir el motor adecuado para el sistema operativo.

Estas situaciones también se producen cuando los estudiantes preparan un examen final, debido al tiempo que suele

pasar entre el fin de cursado del estudiante y el examen.

A partir de esto se generó la idea de crear una plataforma web, accesible a toda hora y desde cualquier sitio con internet, que permita realizar prácticas de SQL y que devuelva un feedback, sobre si es correcta la solución o en caso contrario brinde mensajes claros que indiquen la oportunidad de mejora, colaborando así en el aprendizaje del lenguaje.

Existen plataformas que ofrecen el acceso web, las cuales se describen a continuación y se señalan diferencias:

- Navicat: Software comercial que permite administrar bases de datos de manera remota. El objetivo es netamente comercial y principalmente orientado a Administradores de Bases de Datos. Registrado por PremiumSoft. [12]

- RAT (Relational Algebra Translator): Software que traduce Álgebra Relacional a SQL. El RAT fue desarrollado en la Universidad Nacional de Costa Rica. RAT necesita un software cliente y depende del Sistema Operativo [13].

También se realizaron consultas sobre software que permita mejorar el aprendizaje del lenguaje, pero no fueron encontrados.

Líneas de Investigación, Desarrollo e Innovación

En el GIDTSI una de las líneas de investigación es la Ingeniería en Sistemas de Información y de Software, incluyendo proyectos en calidad, métricas, mantenimiento, trazabilidad, requerimientos, bases de datos, seguridad, relacionados con las áreas de conocimiento: Sistemas de Información, Gestión Ingenieril y Programación. En ese contexto el proyecto desarrolla temáticas relacionadas a las bases de datos, con un importante desarrollo de programación.

En una de las líneas de investigación de las posibles soluciones a la problemática, se pueden realizar aprendizaje a través de ontologías de consultas SQL. Considerando como una ontología [14, 15] al conjunto finito de posibles construcciones SQL que satisfacen el mismo resultado esperado a un ejercicio propuesto. Se podría lograr una ontología para cada ejercicio de práctica, compuesta por los posibles comandos SQL que satisfacen la solución, que el sistema utilizará para validar la solución del usuario. El hecho de poseer los comandos SQL que satisfacen la solución al problema planteado permite guiar al usuario con el conjunto de datos que se deberían obtener en el ejercicio, esta posibilidad de acceder a una vista del resultado esperado (datos), es un vehículo que permite visualizar, a través de ingeniería inversa, cuál debiera ser el comando SQL para llegar al resultado. Tratando las posibles soluciones a un problema de SQL como ontologías, tanto para los comandos, como para el conjunto de datos resultado, el camino correcto sería a través del análisis de patrones, complementado con minería de texto.

El diseño y construcción del prototipo iniciará orientado a la estructura de datos para almacenar el metadato y a una interfaz sencilla para los estudiantes, incluyendo la ejecución de sentencias del sublenguaje DML. Resuelto ello, deberá orientarse al entorno de administración para los docentes.

Resultados y Objetivos

El resultado esperado es que la plataforma permita al usuario:

- Acceder a un intérprete SQL, para ejecutar sentencias SQL, accediendo a una estructura de datos ya poblada.
- Ejecutar sentencias tipificadas, para que los estudiantes elijan.

- Observar el resultado de la ejecución, pudiendo listar y copiar las filas resultantes, cuando es una consulta.
- Recibir mensajes de la plataforma cuando haya un error en la sintaxis.
- Validar el resultado a través de mensajes sobre la cantidad de filas y columnas devueltas por una consulta.
- Recibir mensajes sobre si el resultado es el esperado o qué error se detecta contra el resultado esperado.

El Objetivo General planteado es: Estudiar, investigar, diseñar y proponer una Plataforma Web para colaborar en la capacitación y ejercitación del lenguaje SQL, con bases de datos relacionales. Dicho objetivo se logrará cuando el equipo de investigación logre:

- a) Determinar requerimientos a partir de las necesidades del profesor y del estudiante.
- b) Evaluar las alternativas elegibles, ventajas y desventajas de modelos de persistencia de datos en la web.
- c) Identificar alternativas de integración de diferentes motores de bases de datos.
- d) Apoyar el proceso de selección de lenguajes de programación y tecnología de persistencia de los datos utilizando herramientas multicriterio de evaluación.
- e) Diseñar y desarrollar un prototipo middleware experimental en entorno web.
- f) Gestionar la publicación del middleware con herramientas de integración continua.
- g) Analizar los logs de uso para encontrar relaciones ocultas en el uso de la herramienta. Ejemplo: éxitos y fracasos.
- h) Buscar patrones de comportamiento de los usuarios que permitan definir indicadores de acierto y errores en la escritura.
- i) Construir las ontologías que contribuyan a la sistematización de la observación de los resultados.
- j) Realizar pruebas de validación del esquema propuesto con profesionales.
- k) Medir resultados con el uso de la aplicación en distintos empresas de software de la ciudad de Córdoba.

Formación de Recursos Humanos

El equipo de trabajo está constituido por: un director de proyecto, un codirector, cinco docentes investigadores, dos estudiantes becarios y un graduado.

Seis de los docentes comparten su trabajo en la cátedra de Gestión de Datos, donde se trata la temática de Bases de Datos y específicamente el Lenguaje SQL. El director del proyecto es el director de dicha cátedra, quien desarrollará su tesis de la Maestría en Ingeniería en Sistemas de Información en el seno del proyecto.

El estudiante Agustín Carrasco desarrollará la Práctica Supervisada [16] centrada en la construcción del prototipo y el estudiante Mauricio Spalletti ya ha tenido experiencia como becario en UTN.

Referencias

- [1] Gastañaga, I., Maldonado, C., Martínez Spessot, C. I., & Hintermeister, E. (2006). TecnoDB una Base de Datos Relacional y Prometeo un método de aprovechamiento de Metadatos y Generador de Consultas. In VIII WICC.
- [2] Marciszack, M., Maldonado, C., Martínez Spessot, C. I., Muñoz, R., Navarro, A., Peretti, J. P., & Roggero, L. (2009). Prometeo: una herramienta para el aprovechamiento de metadatos de base de datos relacionales. In XI WICC.
- [3] Vaca, P. A., & Abrutsky, M. (2009). Herramienta para consultas complejas orientada a usuarios finales. In XI WICC.
- [4] Paz Menvielle, M. A., Cuevas, J. C., Damiano, L. E., Muñoz, R., & Quinteros, S. (2013, June). Generador automático de modelos de datos normalizados en bases de datos relacionales. In XV WICC.
- [5] Reinos, E.; Maldonado, C.; Muñoz, R.; Damiano, L.; Abrutsky, M.- Bases de Datos-Edit.AlfaOmega Editores-Argentina - 2012- ISBN: 978-987-1609-31-4.
- [6] Date, C.J.Introducción a los Sistemas de Bases de Datos. Volumen 1. Séptima Edición. México, 2001. Edit. Addison Wesley Longman, Inc. ISBN: 0-201-38590-2.
- [7] Bigus, J. P. Data Mining with Neural Networks: Solving Business Problems from Application Development to Decision Support. 1996. McGraw-Hill. ISBN:0-07-005779-6
- [8] Riquelme, J., Ruiz, R y Gilbert, K. Minería de Datos: Conceptos y Tendencias. Artículo publicado por Revista Iberoamericana de Inteligencia Artificial. ISSN: 1137-3601. <http://www.redalyc.org/articulo.oa?id=92502902>. Último acceso: 11/12/2015, a las 14:35 hs.
- [9] Introducción a los Sistemas de Bases de Datos. Volumen 1. Séptima Edición. México, 2001. Edit. Addison Wesley Longman, Inc. ISBN: 0-201-38590-2
- [10] Elmasri, R. y Navathe, S. Fundamentos de Sistemas de Bases de Datos. 5ta Edición, 2007 Edit. Pearson. ISBN: 978-84-7829-085-7.
- [11] Silberschatz y otros- Fundamentos de Bases de Datos - Quinta Edición- Edit. Mc Graw Hill- Estados Unidos - 2006 - ISBN: 84-481-4644-1.
- [12] <http://www.navicat.com/es/>. Último acceso: 25/02/2016, a las 12:30 hs.
- [13] <http://www.slinfo.una.ac.cr> Último acceso: 20/12/2015, a las 17:45 hs.
- [14] Mario Bunge – Tratado de Filosofía – Ontología I – Editorial Gedisa – Barcelona 2011 – ISBN: 978-84-9784-196-2
- [15] María Jesús Lamarca Lapuente– Ontologías.<http://www.hipertexto.info/documentos/ontologias.htm> . Último acceso: 12/12/2015, a las 17:00 hs.
- [16] UTN -FRC - Prácticas Supervisadas <http://www.institucional.frc.utn.edu.ar/sistemas/Areas/Alumnos/PracticasSupervisadas.asp>: Último acceso: 20/02/2015, a las 10:15 hs.