

# Presentación de Avances en el Desarrollo de una Metodología de Aprovechamiento de Metadatos – Proyecto PROMETEO

Integrantes:

Ing. Cesar Martínez Spessot

Ing. Erica Hintermeister

Ing. Calixto Maldonado

Laboratorio de Investigación de Software  
Departamento de Sistemas  
Facultad Regional de Córdoba  
Universidad Tecnológica Nacional

Dirección Postal  
Universidad Tecnológica Nacional - Facultad Regional Córdoba  
Maestro M. López esq. Cruz Roja Argentina  
Ciudad Universitaria - C.P. (5016)  
Córdoba - República Argentina.

Dirección Electrónica  
cispessot@hotmail.com  
erica@guarani.unc.edu.ar  
calixtomaldonado@hotmail.com

**Abstract.** El presente trabajo reúne los avances en la construcción del sistema generador de consultas Prometeo, presentado en el Simposio Argentino de Sistemas de Información ASIS, de la 34JAIIO2005, presentando las pantallas y estructura del software construido. Contiene una introducción al proyecto, sus antecedentes, e informa finalmente sobre el estado actual del proyecto Prometeo.

## Palabras Clave

Construcción de conocimiento en base a Metadatos, Diccionario de datos, Bases de Datos Relacionales, SQL.

## Introducción

El proyecto Prometeo, presentado en ASIS 2005 con el nombre: “TecnoDB: Desarrollo de una Metodología de Aprovechamiento de Metadatos de los Diccionarios de Datos de Bases de Datos Relacionales para Lograr un Generador de Sentencias SQL – PROMETEO”, fue consolidándose en el avance de la construcción del software, con el propósito de aplicar una Heurística presentada en ese trabajo, y así generar la mayor cantidad de sentencias SQL, basándose en los enunciados de heurística y leyendo los metadatos de los diccionarios de, en una primera etapa, PostgreSQL. Este avance dio los primeros frutos y logró apoyo de la Agencia Córdoba Ciencia con el programa de Apoyo a Grupos de Reciente Formación, estímulo que el grupo agradece. [1]

El envío del presente trabajo a este Simposio responde al Tópico de construcción de conocimiento o ‘Knowledge Discovery and Data Mining’ entendiéndolo que Prometeo hace un análisis de los metadatos y los traduce en asistencia al usuario para construir sentencias sql y construye por sí mismo estas sentencias almacenándolas en la base de datos en un paso posterior de construcción.

## Historia

En el año 2002, dos autores del presente presentaron en CACIC 2002, junto con Anabel Ruiz, un paper sobre el desarrollo de un motor de bases de datos llamado TecnoDB [2], este paper posteriormente fue publicado en el Journal of Science and Technics de la Universidad de la Plata [3]. En el año 2004 TecnoDB logra su acreditación en la Secretaría de Ciencia y Técnica de la Universidad Tecnológica Nacional [4]

Del trabajo nombrado surge la idea de construir un generador de consultas que utilice los metadatos almacenados y de la idea, que aun no podía aplicarse en TecnoDB por el poco desarrollo de su diccionario, se inició este subproyecto para utilizar los metadatos almacenados en diccionarios de datos de motores de bases de datos comerciales. La experiencia de herramientas comerciales que permiten construir sql con Asistentes, llamado en los años 80 como Query-By-Example, inspiraron a los integrantes en el camino

## Objetivo de Prometeo

Durante el proyecto “Desarrollo de un Motor de Base de Datos Relacional, TecnoDB” se dio la percepción sobre la falta de herramientas que asistan al usuario de nuestro motor para escribir sentencias SQL en forma automática, basadas en la información residente en el diccionario de Datos de la Base de Datos objetivo.

Sumado esto a la experiencia con herramientas como Discoverer, Query-Builder de Oracle sirvieron de estímulo y base para diseñar una herramienta que, aparte de hacer a pedido y con intervención del usuario, una sentencia SQL tiene como propósito, con un proceso por lotes o batch recorrer los metadatos de las tablas seleccionadas y las relacionadas, a través de constraints referenciales, escribir automáticamente todas las consultas y a guardar en un esquema simple de tablas propias de la herramienta.

## Presentación de Avances

En el paper presentado en la JAIIO 34 [6] decíamos “Vimos como características necesarias para Prometeo que debería ser **Multiplataforma**, refiriéndonos a multiplataforma a la posibilidad de ser utilizado independientemente del sistema Operativo del equipo cliente...”

Esto motivo en la selección del lenguaje PHP y servidor Apache como plataformas de desarrollo de Prometeo

Para mostrar los avances logrados recurriremos a la presentación citada para nombrar los objetivos detallados, y a continuación de cada uno de ellos, mostrar los avances.

## Muestra de los Avances:

Decíamos en el paper citado:

“En síntesis, que Prometeo pueda generar :

Consultas simples con SELECT,

Consultas multitas de las tablas relacionadas con Constraints Referenciales

DML para insertar, borrar y actualizar columnas

DDL para creación, modificación y borrado de objetos como Vistas ...”

Prometeo actualmente puede cumplir con el primer y cuarto punto:

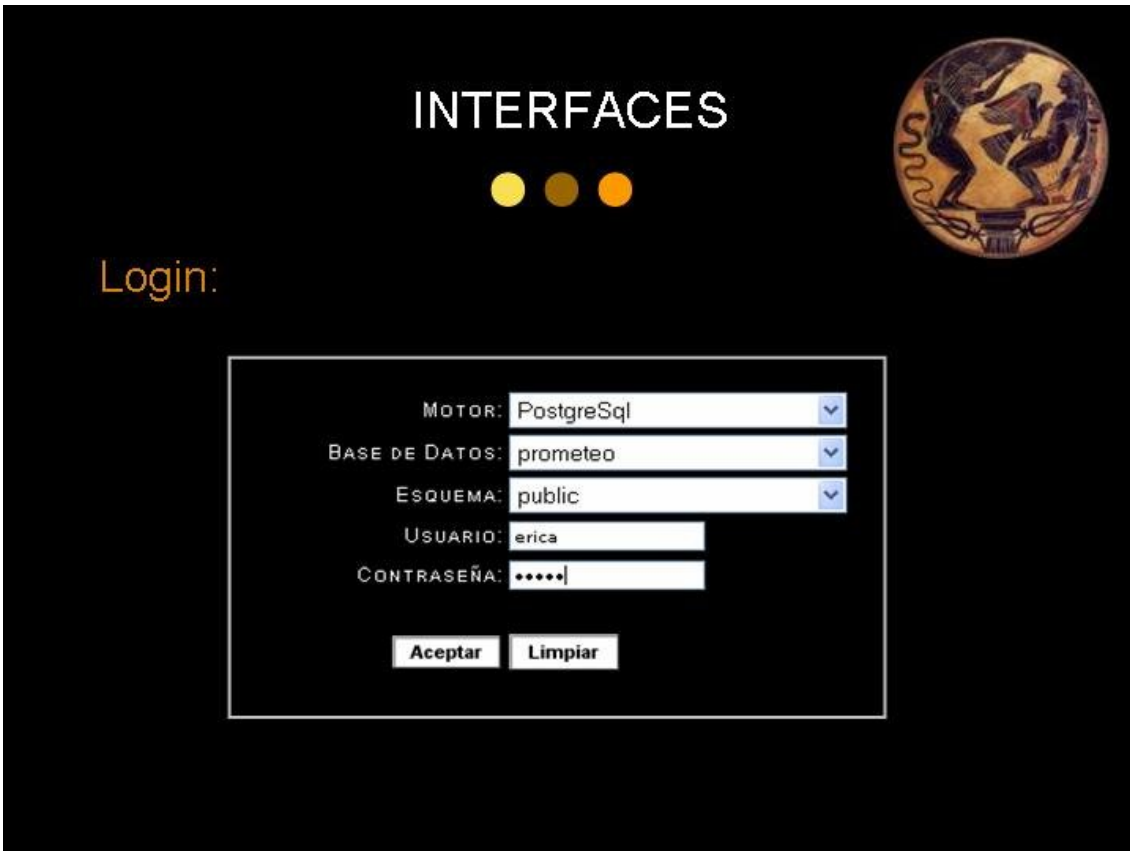
## Comparación De Lo Establecido Y Lo Logrado

*Detalle de las características Iniciales de Prometeo:*

*Objetivo planteado 1*

“Prometeo deberá contactarse con el diccionario de datos de la base de Datos Objetivo y una vez definido el nombre del conjunto de tablas o esquema sobre el que se va a trabajar, creara y guardara, en el modelo de datos propio, todas las consultas posibles generadas. “

Prometeo se conecta con la base elegida, en el ejemplo es la base PostgreSQL:



The screenshot shows a login interface for the Prometeo application. At the top, the word "INTERFACES" is displayed in white capital letters. To the right of this text is a circular logo depicting a figure holding a globe. Below the title are three colored dots (yellow, brown, orange). The word "Login:" is written in orange on the left side. The main form contains the following fields and buttons:

- MOTOR: PostgreSQL (dropdown menu)
- BASE DE DATOS: prometeo (dropdown menu)
- ESQUEMA: public (dropdown menu)
- USUARIO: erica (text input)
- CONTRASEÑA: ..... (password input)
- Buttons: Aceptar, Limpiar

Figura 1. pantalla de conexión a la base PostgreSQL, base de datos Prometeo y esquema publico

*Objetivo planteado 2*

“Consultar tablas, vistas y objetos y sus columnas del diccionario de Datos, en forma general o a través de criterios de a uno, o por esquemas o por nombres de tablas de diferentes esquemas.”



Figura 2: Pantalla de selección de tablas y opciones para hacer la selección de columnas o generar todas las consultas posibles.

A continuación, elegida la tabla, mostrara las columnas para que el usuario elija las integrantes del Select



Figura 3. Pantalla de selección de columnas, si participan del order by y el nombre del alias de columna

```

SELECT PROJECT_ID AS "NROPROYECTO", QUERY_ID AS "NROCONSULTA", LINE_ID AS "NROLINEA", TYPE AS "TIPODATO", TEXT AS
"TEXTOSQL"
FROM LINES
ORDER BY PROJECT_ID ASC,QUERY_ID ASC,LINE_ID ASC,TYPE ASC,TEXT ASC

```

**RESULTADOS DE LA CONSULTA**

NroProyecto	NroConsulta	NroLinea	TipoDato	TextoSql
2	1	1	SELECT	query_id
2	1	2	SELECT	line_id
2	1	3	SELECT	type
2	1	4	SELECT	text
2	1	5	SELECT	project_id
2	1	6	FROM	lines
3	1	7	SELECT	query_id
3	1	8	SELECT	line_id
3	1	9	SELECT	type
3	1	10	SELECT	text

Figura 4: la pantalla devuelve el texto de la consulta generada y el resultado de su ejecución.

*Objetivo planteado 3*

“Crear en base a las consultas anteriores Vistas con el nombre por defecto o darle la posibilidad de darle un nombre distinto y guardar registro de documentación sobre esa vista en las tablas especiales de la Herramienta.”

The screenshot shows a software interface with the following elements:

- Logo: **db Tecnodb**
- Title: **INTERFACES**
- Text: "Selección de vistas creadas según la heurística:"
- Dropdown menu: "VISTAS (5): Seleccione..." with a list of options:
  - Seleccione...
  - productos\_v (circled in red)
  - get\_productos\_descripcion
  - get\_productos\_id\_producto
  - get\_productos\_id\_proveedor
  - get\_productos\_precio
- Button: "Generar Consulta"
- Image: A circular image of a globe with a snake and a figure, possibly a logo or decorative element.

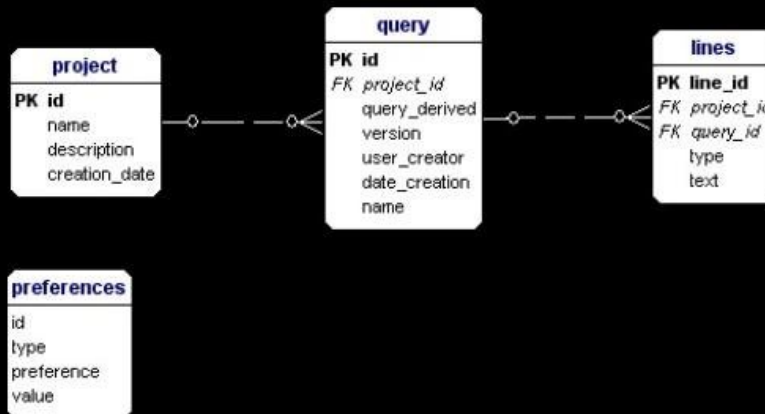
Figura 5: ejemplo de una pantalla que contiene la lista de vistas guardadas en el diccionario destino, con los nombres sugeridos por la Heurística.

*Objetivo planteado 4*

“Las construcciones generadas se guardaran en Proyectos para agruparlos en una entidad de orden superior que permita organizarlos...”

Para estos objetivos deberá contar con modelo de datos que almacene los Metadatos es decir la información sobre las tablas, columnas, constraints o relaciones, índices existentes en la Base de datos. Esto ayudaría al desarrollador, no experto en el modelo de datos, a obtener el resultado sin escribir por sí mismo código SQL.”

# PROMETEO: Modelo de Datos



## Procesos de Prometeo

La figura a continuación refleja la secuencia de pasos que cumple el programa en cumplir los objetivos de acceder a una base de datos, leer el diccionario para construir una consulta, ejecutarla y guardarla.

# PROMETEO: ¿Cómo funciona?



Accede a los metadatos contenidos en el diccionario de datos y detecta:



## Estado Actual

El estado actual esta orientado a la documentación de lo realizado, modularizacion de los procesos orientado a la proxima etapa que es que basados en el avance logrado, acceder a otros diccionarios de Base de datos comerciales como Oracle 10g y a SQL Server 2005.

Continuamos el estudio de nuevas sugerencias de la Heurística definida, que es la base del proyecto para ampliar la utilización de la herramienta.

Faltan aun completar los objetivos restante de los planteados en el paper citado, que es motivo de la actividad en desarrollo.

## Conclusión

Los avances en concretar la idea han sido significativos, pasamos de la idea inicial a un programa que permite extraer algunas sentencias SQL sin escribirlo.

Es aliciente el haber logrado apoyo oficial de la Agencia Córdoba Ciencia, que permitirá costear gastos y viajes para asistir a las presentaciones en Congresos y Jornadas como las JAIIO.

El trabajo que queda por delante es completar los objetivos del paper inicial, y extender su aplicación a nuevas Bases de Datos.

## Referencias

- [1] ACC- Agencia Córdoba Ciencia – Programa de apoyo de grupos de reciente Formacion. –Desarrollo de una metodología de aprovechamiento de Metadatos de los diccionarios de Datos de Bases de Datos Relacionales para lograr un generador de sentencias SQL designado PROMETEO. Ref: 5680 Director: Martinez Spessot, Cesar. <http://www.agenciacordobaciencia.cba.gov.ar/proyectos2005/financiados.htm>
- [2] <http://www.dc.uba.ar/people/congresos/cacic2002/listasaciones1.xls>  
C151- Desarrollo de un Sistema de Gestión de Base de Datos Relacional- Anabel Natalia Ruiz-Cesar Martinez Spessot, Calixto Maldonado
- [3] Journal of Computer Science & Technology 2003, oct, vol.3, no. 2  
Búsqueda por autor en Cesar Spessot, Calixto Maldonado, Anabel Ruiz o por articulo ARG-UNLP-ART-000000256 en <http://www.sedici.unlp.edu.ar/search/search.php>
- [4] TecnoDB: Desarrollo de una Metodología de Aprovechamiento de Metadatos de los Diccionarios de Datos de Bases de Datos Relacionales para Lograr un Generador de Sentencias SQL – PROMETEO  
<http://www.sel.unsl.edu.ar/asis2005/>

## Bibliografia

ACM SIGMOD Record, Volume 29 Issue 1  
SQL standardization: the next steps Andrew Eisenberg, Jim Melton March 2000

Database & data management: The forest and the trees: using oracle and SQL server together to teach ANSI-standard SQL-Gary B. Randolph-October 2003  
Proceeding of the 4th conference on Information technology curriculum

SQLator: an online SQL learning workbench  
Shazia Sadiq, Maria Orłowska, Wasim Sadiq, Joe Lin  
June 2004  
ACM SIGCSE Bulletin , Proceedings of the 9th annual SIGCSE conference on Innovation and

technology in computer science education, Volume 36 Issue 3

Efficient evaluation of XML middle-ware queries  
Mary Fernandez, Atsuyuki Morishima, Dan Suciu  
May 2001

ACM SIGMOD Record , Proceedings of the 2001 ACM SIGMOD international conference on  
Management of data, Volume 30 Issue 2  
An SQL-based query language for networks of relations  
Amit Basu, Rafiul Ahad  
March 1990

ACM SIGMOD Record, Volume 19 Issue 1  
Retrieval effectiveness of an ontology-based model for information selection  
Latifur Khan, Dennis McLeod, Eduard Hovy  
January 2004

The VLDB Journal The International Journal on Very Large Data Bases, Volume 13 Issue 1  
An algebraic approach to static analysis of active database rules  
Elena Baralis, Jennifer Widom  
September 2000

ACM Transactions on Database Systems (TODS), Volume 25 Issue3  
Research sessions: XML I: Storing and querying ordered XML using a relational database  
system  
Igor Tatarinov, Stratis D. Viglas, Kevin Beyer, Jayavel Shanmugasundaram, Eugene Shekita,  
Chun Zhang June 2002  
Proceedings of the 2002 ACM SIGMOD international conference on Management of data