



CONSTRUCCIÓN DE UN DATAMART ACADÉMICO DE LAS PRUEBAS SABER PRO PARA LA INSTITUCIÓN UNIVERSITARIA COLEGIO MAYOR DEL CAUCA

Fredy Alonso Vidal Alegría

**Institución Universitaria Colegio Mayor del Cauca
Popayán, Colombia**

Silvio Ricardo Timarán Pereira

**Universidad de Nariño
Pasto, Colombia**

Resumen

En este artículo se presenta los resultados del proyecto de investigación que tuvo como objetivo la construcción de un datamart académico que almacene los resultados en las competencias genéricas de las pruebas SaberPro que presentaron los estudiantes de la Institución Universitaria Colegio Mayor del Cauca, entre los años 2011 al 2014. La metodología propuesta por Kimball fue utilizada para diseñar el modelo multidimensional del mercado de datos. El datamart se integró a la herramienta de Inteligencia de Negocios libre llamada Pentaho para realizar el análisis multidimensional y la visualización de los resultados. Este datamart brinda información de calidad que ayudará a los directivos de la Institución a tomar decisiones acertadas y oportunas en lo relacionado con la formulación de estrategias para el mejoramiento de la calidad educativa y por ende en el mejoramiento de los resultados en las pruebas SaberPro. Estas pruebas se han convertido en un elemento esencial en la medición del modelo de indicadores de desempeño de la calidad de la educación superior en Colombia.

Palabras clave: DataMart; inteligencia de negocios; SaberPro

Abstract

The results of the research project that had as goal the construction of an academic DataMart that stores the results in the generic skills of the SaberPro tests that presented the students of the Colegio Mayor of Cauca University Institution between the years 2011 to 2014 are presented. The methodology, proposed by Kimball, was used for designing the multidimensional model of the datamart. The datamart was integrated with the free Business Intelligence tool named Pentaho to perform multidimensional analysis and visualization of results. This datamart provides quality information that will help the Institution's executives to make wise and timely decisions regarding the formulation of strategies for improving the quality of education and, therefore, to improve the results of the SaberPro tests. These tests have become an essential element in the Measurement of the model of indicators of the quality of higher education in Colombia.

Keywords: business intelligence; DataMart; SaberPro

1. Introducción

El Decreto 3963 de octubre 14 de 2009, reglamentó el examen de Estado de calidad de educación superior SaberPro, como un instrumento periódico y estandarizado para la evaluación externa de la calidad de la educación superior, establece que son objeto de evaluación las competencias de los estudiantes próximos a graduarse en los distintos programas de pregrado. El examen está compuesto por pruebas que evalúan competencias genéricas y específicas. De acuerdo a los lineamientos Saber Pro del Instituto Colombiano para la Evaluación de la Educación (ICFES), todos los estudiantes deben presentar los módulos de competencias genéricas sin importar el programa de formación que cursen, que incluye competencias de lectura crítica, razonamiento cuantitativo, comunicación escrita, inglés y competencias ciudadanas (ICFES, 2011).

Dentro de los objetivos de las pruebas SaberPro están (ICFES, 2016):

- Comprobar el grado de desarrollo de las competencias de los estudiantes próximos a culminar los programas académicos de pregrado que ofrecen las instituciones de educación superior.
- Producir indicadores de valor agregado de la educación superior en relación con el nivel de competencias de quienes ingresan a este nivel; proporcionar información para la comparación entre programas, instituciones y metodologías, y mostrar su evolución en el tiempo.
- Servir de fuente de información para la construcción de indicadores de evaluación de la calidad de los programas e instituciones de educación superior y del servicio público educativo, que fomenten la cualificación de los procesos institucionales y la formulación de políticas, y soporten el proceso de toma de decisiones en todos los órdenes y componentes del sistema educativo.

En este artículo se presenta los resultados del proyecto de investigación que tuvo como objetivo la construcción de un datamart académico que almacene los resultados en las competencias genéricas de las pruebas SaberPro que presentaron los estudiantes de la Institución Universitaria Colegio Mayor del Cauca, entre los años 2011 al 2014 con el fin de que brinde información de calidad que ayude a los directivos de la Institución a tomar decisiones acertadas y oportunas en lo relacionado con la formulación de estrategias para el mejoramiento de la calidad educativa de los programas que ofrece y por ende en el mejoramiento de los resultados en las pruebas SaberPro, que se han convertido en un elemento esencial en la medición del modelo de indicadores de desempeño de la calidad de la educación superior en Colombia.

2. Metodología

Para el desarrollo del datamart académico de las pruebas SaberPro 2011 – 2014 para la Institución Universitaria Colegio Mayor Del Cauca, se utilizó la metodología propuesta por Kimball (1998), la cual se basa en lo que se denomina Ciclo de Vida Dimensional del Negocio (*Business Dimensional Lifecycle*) que contempla las siguientes etapas:

Planificación del Proyecto: En este proceso se determina el propósito del proyecto de la Bodega de Datos, sus objetivos específicos y el alcance del mismo, los principales riesgos y una aproximación inicial a las necesidades de información.

Definición de Requerimientos del Negocio: La definición de requerimientos, es un proceso de entrevistar al personal de negocio y técnico, aunque siempre conviene, tener un poco de preparación previa. En esta tarea, se debe aprender sobre el negocio, los competidores, la industria y los clientes del mismo.

Modelado Dimensional: Es un proceso dinámico y altamente iterativo. Comienza con un modelo dimensional de alto nivel obtenido a partir de los procesos priorizados y descritos en la definición de Requerimientos del Negocio.

Diseño e Implementación del ETL: Se diseña e implementa el sistema de Extracción, Transformación y Carga (ETL) de datos, en el cual se soporta la Bodega de Datos.

Implementación y Análisis de Resultados: en esta fase se establecerá el resultado final del proyecto, su ordenamiento y presentación al cliente y como se puede reincorporar a la información que ya se cuenta. En este capítulo se confronta los resultados obtenidos con la implementación del proyecto.

3. Resultados del Proyecto

En la fase de *planificación de proyecto* se elaboró el acta de constitución del proyecto (*Project Charter*), en la cual se detallan cada uno de los aspectos fundamentales y cruciales del proyecto, es aquí donde se delimitó el alcance, se definieron los objetivos, los entregables, las partes interesadas (*Stakeholder*, Clientes), se asignaron

responsabilidades y las consideraciones (Riesgos, asunciones, restricciones) del proyecto. En la figura 1 se muestra el acta de constitución del proyecto.

En la fase de *definición de requerimientos* se realizaron entrevistas a las partes interesadas, se identificaron los temas analíticos y procesos de negocio. A partir de ahí se procedió a elaborar la matriz de procesos/dimensiones (*Bus Matrix*), esta matriz tiene en sus filas los procesos del negocio identificados, y en las columnas, las dimensiones identificadas. En la tabla 1 se puede observar la matriz de procesos/dimensiones del proyecto:

Project Charter

GERENTE	FREDY ALONSO VIDAL A.	DOCUMENTO PREPARADO POR	FREDY ALONSO VIDAL JUAN CARLOS ALEGRIA.
NOMBRE DEL PROYECTO	CONSTRUCCION DE UN DATAMART ACADEMICO DE LAS PRUEBAS SABER PRO 2011 – 2014 PARA LA INSTITUCIÓN UNIVERSITARIA COLEGIO MAYOR DEL CAUCA		FECHA JULIO 29 DE 2016

Resumen:

En el presente documento se presenta la propuesta de construcción de un Data ~~Mart~~ académico que almacene los resultados de las pruebas saber pro en las vigencias 2011, 2012, 2013 y 2014 presentadas por los estudiantes de la Institución Universitaria Colegio Mayor del Cauca; soportado en un modelamiento dimensional que facilite la toma de decisiones pertinentes para mejorar los resultados de dichas pruebas, debido a que las pruebas SABER PRO son una dimensión esencial en la medición del Modelo de Indicadores de Desempeño de la Educación Superior, en los cuales la Institución no ha sido bien calificada, afectando su imagen y reconocimiento institucional, incluso dificultando la consecución de recursos financieros a nivel nacional.

Gerente del Proyecto Asignado:

Fredy Alonso Vidal Alegria, estudiante de la Especialización en Administración de la Información y Bases de Datos que se orienta en la Institución Universitaria Colegio Mayor del Cauca en el primer periodo del año 2016 y docente de la Institución.

1.01_ProjectCharter (Carta del Proyecto)

Propósito / Necesidades a suplir / Consideraciones Estratégicas:
OBJETIVO GENERAL - Construir un Data Mart Académico de las pruebas saber pro 2011 – 2014 para la Institución Universitaria Colegio Mayor del Cauca, con el fin de realizar el análisis de resultados de dichas pruebas para la toma de decisiones pertinentes relacionadas con el mejoramiento de la calidad de los programas y el posicionamiento institucional. OBJETIVOS ESPECÍFICOS - Realizar la exploración y análisis de calidad de los datos de las pruebas saber Pro en las vigencias 2011 - 2014. - Realizar los procesos de Extracción, Transformación y Almacenamiento de los datos al Data Mart. - Diseñar e implementar el modelo multidimensional del DataMart, en un Sistema de Bases de Datos Relacional. - Integrar el Data Mart con una herramienta de Inteligencia de negocios libre para la visualización de los datos. - Evaluar los resultados obtenidos para la toma de decisiones eficaces en la relacionado a las pruebas Saber Pro en la CMC.
Descripción del Producto y Entregables:
Para el desarrollo Data Mart Académico de las pruebas saber pro 2011 – 2014 para la Institución Universitaria Colegio Mayor Del Cauca, se siguen los pasos definidos por la metodología Kimball: - Planificación del Proyecto: En este proceso se determina el propósito del proyecto de DW/BI, sus objetivos específicos y el alcance del mismo, los principales riesgos y una aproximación inicial a las necesidades de información. - Definición de Requerimientos del Negocio: La definición de requerimientos, es un proceso de entrevistar al personal de negocio y técnico, aunque siempre conviene, tener un poco de preparación previa. En esta tarea, se debe aprender sobre el negocio, los competidores, la industria y los clientes del mismo. - Modelado Dimensional: Es un proceso dinámico y altamente iterativo. Comienza con un modelo dimensional de alto nivel obtenido a partir de los procesos priorizados y descritos en la definición de Requerimientos del Negocio.

1.01_ProjectCharter (Carta del Proyecto)

• Diseño e Implementación del proceso de Extracción, Transformación y Carga (ETL) de datos, en el cual se soporta el Data Warehouse. • Análisis de Resultados: en esta fase se establecerá el resultado final del proyecto, su ordenamiento y presentación al cliente y como se puede reincorporar a la información que ya se cuenta. En este capítulo se confronta los resultados obtenidos con la implementación del proyecto.		
Supuestos y Restricciones:		
La Institución Universitaria Colegio Mayor del Cauca, asume los costos financieros de los recursos humanos, gerente del proyecto. El tiempo máximo de ejecución del proyecto es de 4 meses que equivale a la duración de un periodo académico.		
Riesgos identificados:		
• El Colegio Mayor del Cauca, no asume los costos del recurso humano del proyecto, por lo que se debería buscar una nueva fuente de financiación y el proyecto no se puede llevar a cabo. • Problemas con la calidad de los datos existentes en los datos entregados por el ICFES. • Poca disponibilidad del talento humano experto en el área de la inteligencia de negocios.		
Presupuesto Estimado:		
Para lograr el proyecto, se tiene presupuestado un costo de \$ 27.073.628,00 , distribuidos de la siguiente manera:		
Recursos Humanos \$21.513.728,00		
Recursos Hardware \$ 4.470.000,00		
Recursos Software \$ 719.900,00		
Materiales e Insumos \$ 370.000,00		
TOTAL \$ 27.073.628,00		
Autorización de proyecto:		
<table border="1"> <tr> <td> Firma del Gerente del Proyecto</td> <td> Firma Rector</td> </tr> </table>	 Firma del Gerente del Proyecto	 Firma Rector
 Firma del Gerente del Proyecto	 Firma Rector	

1.01_ProjectCharter (Carta del Proyecto)

Figura 1. Acta de Constitución del Proyecto

Tabla 1. Matriz de procesos/dimensiones (Bus Matrix)

PROCESOS DEL NEGOCIO	DIMENSIONES					
	Estudiantes	Programas	Pruebas (Comp. Ciudadanas, Lectura Crítica, Comunicación Escrita, Razonamiento Cuantitativo e Ingles	Lugares	Instituciones	Tiempo
Promedio general de algunas instituciones del municipio de Popayán	X	X	X	X	X	X
Promedio general de todos los programas de la IUCMC correspondientes a los años 2011 a 2014.	X	X	X	X	X	X

Media general por programa académico, por tipo de prueba y por año.	X	X	X	X	X	X
---	---	---	---	---	---	---

En la fase del *modelado dimensional* se identificaron el nivel de granularidad, las dimensiones, las medidas y las tablas de hechos del datamart. Para obtener este modelo dimensional se realizó un gráfico denominado modelo dimensional de alto nivel o gráfico de burbujas (*Bubble chart*), tal como lo ilustra en la figura 2.

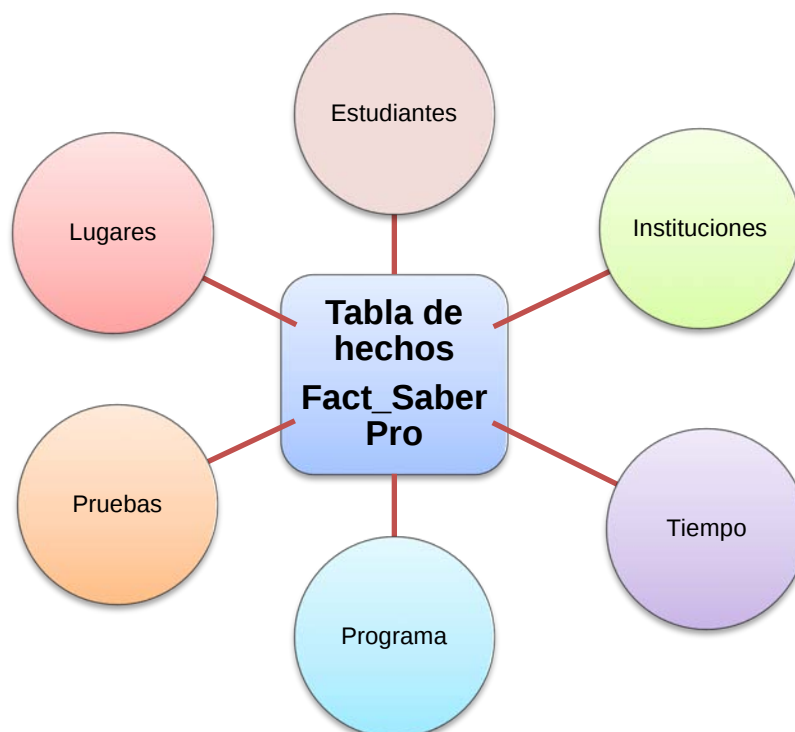


Figura 2. Modelo final de alto nivel de la sesión inicial de diseño del datamart

A partir del modelo dimensional de alto nivel se diseñó el modelo dimensional en estrella del datamart y se implementó en el sistema gestor de bases de datos PostgreSQL. La topología en estrella se muestra en la figura 3.

En la fase del *diseño e implementación del sistema de Extracción, Transformación y Carga (ETL)* de datos, que permite almacenar datos limpios y transformados al datamart académico, se diseñaron una serie de scripts en el lenguaje SQL que permitieron limpiar y transformar los datos de las bases de datos de las pruebas SaberPro del ICFES (fuentes de datos iniciales) y cargarlos al datamart. Para automatizar este proceso, se desarrolló una aplicación bajo Delphi™, con opciones para realizar reglas de

integración, análisis de calidad de datos limpieza, discretización y eliminación de atributos.

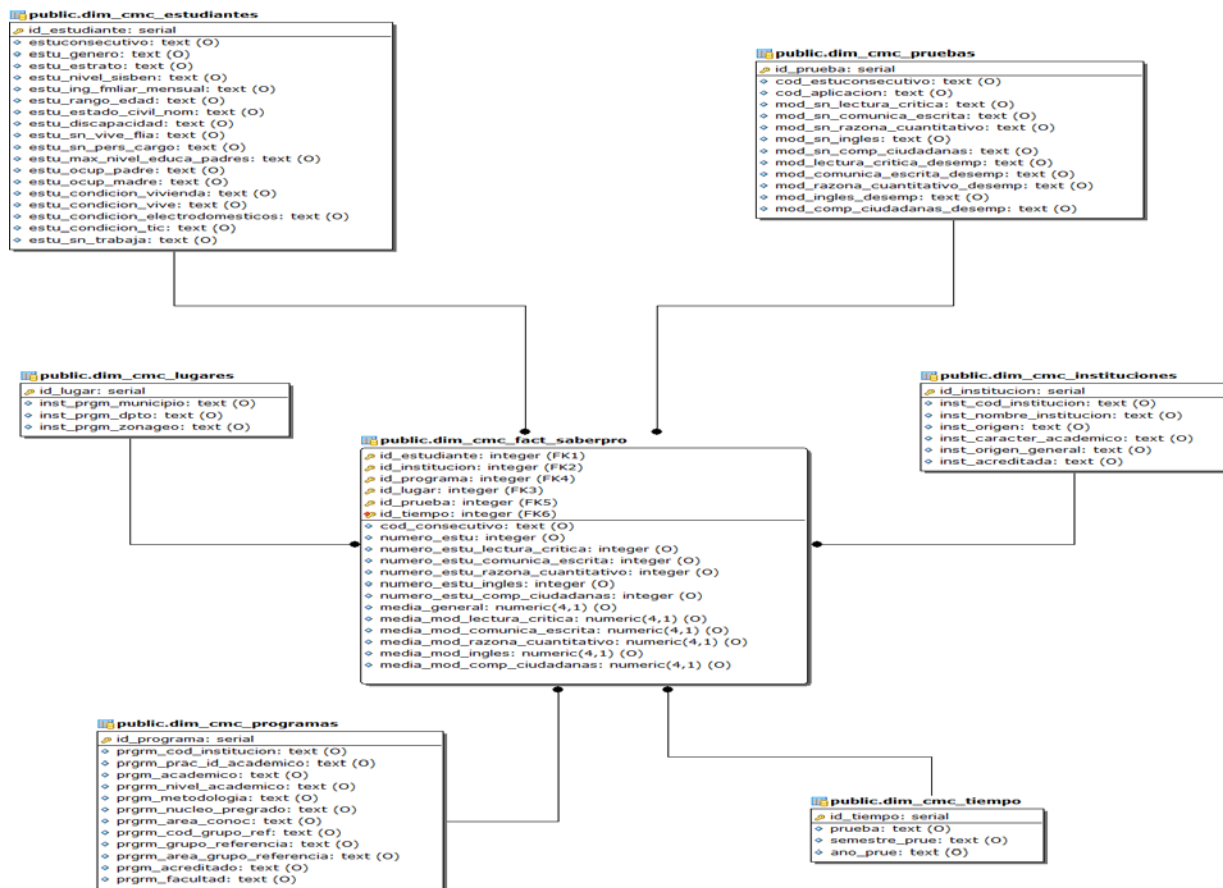


Figura 3. Modelo dimensional en estrella del datamart

En la fase de *implementación y análisis de resultados*, se utilizó la herramienta BI Server Pentaho™, para construir los cubos a partir del modelo dimensional del datamart, construido en PostgreSQL. Además, para generar un sinnúmero de reportes y facilitar la presentación, interpretación y visualización de los datos se utilizó la herramienta Pentaho Report Designer™ (PRD). De este modo, el análisis de los resultados obtenidos permitió conocer los siguientes aspectos:

- El Colegio Mayor del Cauca, se consolida como la segunda institución de educación superior en Popayán después de la Universidad del Cauca en resultados de las pruebas SaberPro.
- El programa Tecnología en Diseño Artesanal del Colegio Mayor del Cauca, presenta el mejor promedio en las pruebas Saber Pro durante las vigencias 2011 - 2014, seguido de la Tecnología en Desarrollo de Software y el programa Profesional en Ingeniería Informática.

- El rendimiento histórico más alto para el área de inglés se obtuvo en el programa de Tecnología en Diseño Artesanal en el periodo 2 de 2011. Así mismo, el rendimiento más bajo se obtuvo en el programa de Tecnología en Delineantes de Arquitectura en el periodo 1 de 2012. De igual forma es importante mencionar que el programa con mejor promedio a lo largo de los años es el de Ingeniería Informática con un valor de 10.355, mientras que el programa con menor promedio es el de Tecnología en Delineantes de Arquitectura con un valor de 9.55. Finalmente, los programas que presentan un comportamiento más parejo a lo largo del tiempo son: Administración Financiera, Arquitectura, Ingeniería Informática.
- El rendimiento histórico más alto para el área de Competencia Ciudadana se obtuvo en el programa de Tecnología en Desarrollo de Software en el periodo 2 de 2012. Así mismo, el rendimiento más bajo se obtuvo en el programa de Tecnología en Delineantes de Arquitectura en el periodo 1 de 2012. De igual forma es importante mencionar que el programa con mejor promedio a lo largo de los años es el de Tecnología en Finanzas y Auditoría Sistematizada con un valor de 9.885, mientras que el programa con menor promedio es el de Tecnología en Delineantes de Arquitectura con un valor de 9.365. Finalmente, los programas que presentan un comportamiento más parejo a lo largo del tiempo son: Administración Financiera e Ingeniería Informática.
- El rendimiento histórico más alto para el área de Comunicación Escrita se obtuvo en el programa de Tecnología en Diseño Artesanal en el periodo 2 de 2012. Así mismo, el rendimiento más bajo se obtuvo en el programa de Tecnología en Delineantes de Arquitectura en el periodo 2 de 2012. De igual forma es importante mencionar que el programa con mejor promedio a lo largo de los años es el de Tecnología en Diseño Artesanal con un valor de 10.23, mientras que el programa con menor promedio es el de Tecnología en Delineantes de Arquitectura con un valor de 9.16. Finalmente, los programas que presentan un comportamiento más parejo a lo largo del tiempo son: Administración Financiera, Arquitectura y Tecnología en Delineantes de Arquitectura.
- El rendimiento histórico más alto para el área de Lectura Crítica se obtuvo en los programas de Tecnología en Desarrollo de Software y Tecnología en Gestión Empresarial para el periodo 2 de 2012. Así mismo, el rendimiento más bajo se obtuvo en el programa de Tecnología en Delineantes de Arquitectura en el periodo 1 de 2012. De igual forma es importante mencionar que el programa con mejor promedio a lo largo de los años es el de Tecnología en Desarrollo de Software con un valor de 10.01, mientras que el programa con menor promedio es el de Administración Financiera con un valor de 9.6. Finalmente, los programas que presentan un comportamiento más parejo a lo largo del tiempo son: Administración de Empresas y Tecnología en Gestión Comercial y de Mercados.

- El rendimiento histórico más alto para el área de Razonamiento Cuantitativo se obtuvo en los programas de Tecnología en Diseño Artesanal para el periodo 2 de 2012. Así mismo, el rendimiento más bajo se obtuvo en el programa de Tecnología en Diseño Artesanal en el periodo 2 de 2011. De igual forma es importante mencionar que el programa con mejor promedio a lo largo de los años es el de Ingeniería Informática con un valor de 10.21, mientras que el programa con menor promedio es el de Administración Financiera con un valor de 9.53. Finalmente, el programa que presenta un comportamiento más parejo a lo largo del tiempo es el de Ingeniería Informática.

4. Conclusiones y Trabajos Futuros

A modo de conclusión es importante mencionar que la Institución Universitaria Colegio Mayor del Cauca cuenta con un datamart académico que almacena datos históricos, limpios y transformados sobre los resultados en las competencias genéricas de las pruebas SaberPro que presentan los estudiantes de sus diferentes programas. Con este datamart integrado a la herramienta Pentaho los directivos de la institución podrán realizar fácilmente diferentes tipos de análisis para obtener información de calidad y producir indicadores que les permitan compararse entre programas, instituciones y metodologías, y mostrar su evolución en el tiempo en lo relacionado a las pruebas SaberPro, con el fin de mejorar cada vez la calidad de la educación que se imparte en sus programas.

Como trabajos futuros están el capacitar a los funcionarios y directivos de la Institución Universitaria Colegio Mayor del Cauca para que puedan utilizar esta herramienta de Inteligencia de Negocios fácilmente y actualizar el datamart académico con los resultados de las nuevas pruebas SaberPro de cada año.

5. Referencias

- ICFES (2011), Lineamientos Saber Pro. Instituto Colombiano para la Evaluación de la Educación. Bogotá, Colombia. Consultado el 20 de septiembre de 2016 en http://aprendeenlinea.udea.edu.co/lms/moodle/file.php/532/Lineamientos_SABER_PRO_2011_2_30_08_1_.pdf.
- ICFES. (2016). Guías de evaluación. Competencias Genéricas. Instituto Colombiano para la Evaluación de la Educación. Bogotá, Colombia. Consultado el 20 de septiembre de 2016 en <http://www.icfes.gov.co>.
- Kimball, R., Reeves, L., Ross, M. and Thomhwaite, W. (1998). The Data Warehouse Lifecycle Toolkit. Ed. John Wiley & Sons. New York, USA.

Sobre los autores

- **Fredy Alonso Vidal Alegría:** Ingeniero de Sistemas, Especialista en Administración de la Información y Bases de Datos. Magister en Educación. Decano de la Facultad de Ingeniería de la Institución Universitaria Colegio Mayor del Cauca. decing@unimayor.edu.co
- **Silvio Ricardo Timarán Pereira:** Ingeniero de Sistemas y Computación, Master of Science en Ingeniería, Doctor en Ingeniería énfasis Ciencias de la Computación. Profesor Titular del departamento de Sistemas de la Facultad de Ingeniería de la Universidad de Nariño. ritimar@udenar.edu.co

Los puntos de vista expresados en este artículo no reflejan necesariamente la opinión de la Asociación Colombiana de Facultades de Ingeniería.

Copyright © 2017 Asociación Colombiana de Facultades de Ingeniería (ACOFI)