



XXIII Comunicaciones Científicas y Tecnológicas

Orden Poster: CT-008 (ID: 644)

Autor: Lopez, Mariano Emanuel Alejandro

Título: Modelo de proceso para proyectos de explotación de información en el ámbito académico

Director:

Palabras clave: Modelo de proceso, Explotación de información, Analítica académica

Área de Beca: Tecnologías

Tipo Beca: Cyt - Pregrado

Periodo: 01/03/2016 al 01/03/2017

Lugar de trabajo: Facultad De Cs. Exactas Y Naturales Y Agrimensura

Proyecto: (13F010) Métodos y herramientas para la calidad del software.

Resumen:

En la actualidad, los volúmenes de datos generados y almacenados por las organizaciones crecen de forma desmesurada. Esta situación trae consigo grandes desafíos, entre ellos cómo integrar y transformar estos datos en información relevante de forma que se pueda optimizar el proceso de toma de decisiones en los negocios. Las instituciones de Educación Superior requieren también tomar decisiones para cumplir sus "objetivos de negocio", que son maximizar la "rentabilidad académica", en términos de lograr la mayor cantidad y calidad de profesionales formados en distintos campos del conocimiento. Para ello necesita identificar, comprender y predecir posibles problemas en el proceso formativo, desde el ingreso hasta la finalización de los estudios.

En este trabajo se propone aprovechar la información disponible en el sistema de gestión de alumnos de la Facultad de Ciencias Exactas y Naturales y Agrimensura de la UNNE, denominado "SIU Guaraní", para extraer valor para un conjunto de indicadores que permitan un monitoreo permanente y confiable de los procesos educativos y sus resultados, a fin de anticipar situaciones, proponer mejoras y apoyar políticas o mecanismos enfocados en mitigar la deserción, el desgranamiento, la baja tasa de graduación, entre otros problemas.

Para lograr el objetivo se propuso diseñar y desarrollar un "Almacén de datos" o DW (Data Warehouse), haciendo énfasis en el modelo de proceso que se requiere para esta aplicación particular en el contexto académico, basado en los conceptos y técnicas correspondientes al campo de la explotación de información en la disciplina Ingeniería del Software.

Para la construcción de la solución DW se seleccionó la metodología "ciclo de vida del almacén de datos" (DW Lifecycle) de Ralph Kimball, la cual consta de tres etapas principales: Planificación del proyecto, Definición de requerimientos del negocio y Administración del Proyecto de DW. Esta última, a su vez, consta de las sub-etapas: Modelado dimensional, Diseño físico, Diseño e implementación del subsistema de ETL e Implementación.

Como resultado de las etapas de la metodología se obtuvo:

Planificación del proyecto: Objetivo confeccionar un modelo de proceso de explotación de información que brinde soporte a la toma de decisiones en el contexto académico y administrativo de la carrera Licenciatura en Sistemas de Información de la FaCENA – UNNE plan 2009, para ello se utilizará el módulo de gestión académica utilizado por la FaCENA– UNNE SG (SIU-Guaraní) como fuente principal de datos y al sistema básico de indicadores de INFOACES (Sistema Integral de Información sobre la Educación Superior en América Latina) y el anuario de estadísticas universitarias de la SPU (Secretaría de Política Universitarias de la República Argentina) como guía para la selección de indicadores.

Definición de requerimientos del negocio: Identificación de los diferentes procesos administrativos y académicos vinculados con la actividad de los alumnos que se gestiona mediante sistema SG, la dependencia entre procesos, la periodicidad de los procesos, las diferentes condiciones que el alumno posee en el tránsito de su carrera de acuerdo a la normativa que regula la permanencia y el avance en la universidad. Definición del conjunto de indicadores en forma conjunta con las autoridades y usuarios directos de la información para el seguimiento de los resultados académicos, categorizados en las siguientes dimensiones: "Ingreso", "Cursado" y "Finalización".

Modelado dimensional: Definición del modelo conceptual del almacén de datos en base a la fuente de datos e indicadores previamente seleccionados, el mismo cuenta con cinco tablas de hechos "datos métricos a analizar" (rendimiento académico, historia académica, reinscripciones, egreso y aspirantes) y cuatro tablas de dimensiones "datos descriptivos" (carrera, materia, alumno y fecha).

Diseño físico: Implementación del modelo conceptual SQL Server.

Diseño e implementación del subsistema de ETL: Scripts en Pentaho Data Integración, los cuales realizan la extracción de los datos seleccionados, limpieza y corrección de los mismos “en los casos necesarios”, y su carga inicial y/o refresco del almacén de datos de forma automatizada.

Desarrollo de aplicaciones de Business Intelligence: Cinco cubos OLAP (On-Line Analytical Processing) generados a partir de las tablas de hechos, los cuales permiten al usuario contar con autonomía alta flexibilidad en las consultas. Y una serie de informes estándar de formato predefinido son generados a partir de los cubos, pueden ser accedidos vía web por usuarios autorizados y ser descargados en los formatos: XML, PDF, CSV, Excel y Word.

Actualmente el proyecto se encuentra en etapa de “Implementación” para su despliegue en el ambiente real de desempeño.