

Nuevas dimensiones y métricas en la información para la toma de decisiones: Aplicación Data WareHouse en Instituciones Universitarias

New dimensions and metrics information for decision -making: Application Data Warehouse on University Institutions

Claudia Minnaard

Instituto de Investigaciones de Tecnologías y Educación
Universidad Nacional de Lomas de Zamora, Argentina
minaard@uolsinectis.com.ar

Diego Servetto

Instituto de Investigaciones de Tecnologías y Educación
Universidad Nacional de Lomas de Zamora, Argentina
diegoservetto@gmail.com

Guadalupe Pascal

Instituto de Investigaciones de Tecnologías y Educación
Universidad Nacional de Lomas de Zamora, Argentina
Guadalupe.pascal@hotmail.com

Unelén Lobo Mirasson

Instituto de Investigaciones de Tecnologías y Educación
Universidad Nacional de Lomas de Zamora, Argentina
uneleen@gmail.com

Resumen

El presente artículo continúa la línea de investigación presentada en el congreso de Workshop de Investigadores en Ciencias de la Computación, el cual hace referencia a la implementación de un sistema Data WareHouse para la toma de decisiones por parte de las autoridades de la Institución Educativa. Anteriormente se mencionó el contexto actual de la Facultad de Ingeniería de la UNLZ respecto a las carencias de información concreta en temas relevantes de la actividad académica. En el presente artículo se busca abordar la segunda etapa de incorporación de nuevas métricas a través de distintos orígenes, orientado a la actividad dentro del Campus Virtual y datos del Cuerpo Docente.

Abstract

This article continues the line of research presented at the conference Researchers Workshop in Computer Science , which refers to the implementation of a data warehouse system for decision-making by the authorities of the educational institution . Previously the current context of the Faculty of Engineering of the UNLZ regarding deficiencies of specific information on relevant topics of academic activity mentioned . In this article it seeks to address the second step of incorporating new metrics across different backgrounds , oriented activity within the Virtual Campus and Faculty data .

Palabras clave / key words: Data WareHouse, Instituciones Universitarias, LCMS, Software de Gestión Académica / Data Warehouse , academic institutions , LCMS, Academic Management Software

1 Introducción

Entendiendo que la Universidad es un ámbito cada vez más complejo, es necesario contar con un sistema fehaciente de información que tenga la capacidad de brindar soluciones rápidas y eficaces, de manera tal que la información esté disponible de manera sencilla y poder contribuir a eliminar la incertidumbre ante la presentación de un problema, por parte de las autoridades. Por este motivo, es que se pretende emplear la tecnología como herramienta de planeación estratégica en el ámbito educativo, orientando el manejo de la información y la comunicación a mejorar el desempeño de las distintas áreas involucradas y por lo tanto, la calidad educativa.

El presente trabajo se encuentra enmarcado en la segunda etapa de la implementación de un sistema Data WareHouse para la toma de decisiones dentro de la Institución Universitaria. Inicialmente se incorporaron métricas referentes al rendimiento académico por cátedra como ser: cantidad de alumnos inscriptos, aprobados, desaprobados, ausentes y accesos al Campus Virtual. En esta instancia se pretende abordar la segunda etapa de incorporación de métricas y dimensiones. A lo largo del artículo, se desarrollarán los objetivos de estas incorporaciones, las técnicas utilizadas para la recopilación de la información, la forma de implementación y el diseño que se ha utilizado.

1.1 Aplicación de la Tecnología

El sistema de Data Warehouse (DW), como su nombre lo indica, hace referencia a un almacén de datos. Estos pueden provenir de distintas fuentes y deben ser estructurados de determinada manera que no haya inconsistencias, asimismo deben tratarse de datos que posean cierta relación entre sí. Por otro lado, ya que estos datos van a representar una tendencia, deben tener un contenido histórico considerable. Como última característica, los datos en este sistema no deben ser volátiles, por lo que se entiende que deben ser permanentes y no modificados. (William Inmon, 2000).

En resumen, lo que ofrece un sistema de DW es poder confeccionar una base de datos fiable, con tablas dinámicas y multivariadas, en donde se pueden convertir datos, entendidos como “un conjunto de hechos aislados y objetivos, relativos a eventos que solo son descritos como registros estructurales de una transacción”, en información, “conjunto de datos ya procesados y que tienen un significado y que por lo tanto tienen utilidad para quienes tienen que tomar decisiones”. (Davenport y Prusak, 1998).

Debido a la complejidad en el intercambio de información en organizaciones como las instituciones universitarias, es necesario contar con un sistema en donde los datos se encuentren disponibles de manera tal que proporcionen información acerca de distintas áreas de interés, en este caso, entorno al rendimiento académico. En una primera etapa, se han incorporado métricas por cada cátedra en donde se pueden visualizar fácilmente la cantidad de alumnos inscriptos, aprobados, desaprobados y ausentes por carrera, ciclo, año académico, cuatrimestre y cátedra. En la misma línea, se incorporaron porcentajes de accesos al Campus Virtual y sus respectivas frecuencias. Es importante destacar que los datos obtenidos respecto de los datos de los alumnos, son relevados del sistema SIU-GUARANÍ, sistema implementado en el año 2005 en la Facultad de Ingeniería, por lo que se cuenta con una data histórica de diez (10) años hasta la actualidad. En el caso de los datos del Campus Virtual, también datan del mismo año y se relevan desde la plataforma E-educativa. En esta etapa de segunda incorporación de datos, se decidió incorporar las métricas del uso del Campus Virtual y la dimensión Docentes.

1.2 Objetivos

En los últimos años, la tecnología en la educación se ha convertido en una herramienta fundamental para la comunicación entre docentes y alumnos, ya que permite una cierta flexibilidad para “atender demandas educativas insatisfechas por la educación convencional” (Portal Educativo del Estado Argentino), en donde existe una interacción entre todos los intervinientes aún más allá de clases asistenciales. A demás, Fernández, Álvarez y Suárez (2006) postulan que la incorporación de las nuevas tecnologías capta la atención de los alumnos generando motivación, lo que se reconoce como un motor en el aprendizaje. Por lo tanto saber cómo el alumno utiliza esta valiosa herramienta puede reflejar en parte su rendimiento dentro del ámbito académico. Se pretende observar cómo a partir de la participación en distintas secciones dentro del Campus Virtual, como foros, descarga de archivos, envío de e-mails internos, noticas y sitios afectan e impactan en la manera en que el alumno se desenvuelve en cada materia.

En relación a la incorporación de la dimensión “Docentes”, la misma, parte de la necesidad de identificar al responsable de cada cátedra y los cambios surgidos a través de los años (entre ellos la incorporación y adaptación al Campus Virtual). La importancia de esta dimensión esta asociada a que el empleo de las TIC en el sistema educativo requiere de una transformación del proceso enseñanza-aprendizaje, tanto en la metodología como en la formación de los docentes (Izquierdo y Pardo, 2007). Para Hamidian (2010), la incorporación de las tecnologías en el sector educativo implica, reconocer y capacitar a los docentes a fin de transmitir los conocimientos de una manera diferente y adaptada. En este caso, los datos se encontraban de manera muy dispersa y el proceso para su recopilación era muy complejo, por lo que además se vio la necesidad de agrupar los datos en una sola base, de manera que su disponibilidad sea más sencilla.

2 Contexto Actual

En los tiempos que corren, la información constituye una parte fundamental en una organización, entendida como “formaciones complejas y plurales compuestas por individuos y grupos que constituyen un sistema de roles coordinados mediante un sistema de autoridad y comunicación articulado por un sistema de significados compartidos” (Díaz de Quijano, 2000).

Actualmente para estudiar el comportamiento de la organización comprendida por la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora, se cuenta con indicadores de Inscriptos/Aprobados/Desaprobados/Ausentes por materia de Ingeniería Industrial e Ingeniería Mecánica, como así también cantidades de Ingresantes por Año/Carrera/Especialidad. En la siguiente figura se representan datos de Inscriptos a Cursadas y porcentaje de Regularización de las carreras en Ing. Industrial y Mecánica con sus dos orientaciones.

Por Carrera	Inscriptos Cursada	% Regularización sin Ausentes
INGENIERIA INDUSTRIAL ORIENTACION GESTION	17,932	84.65 %
INGENIERIA INDUSTRIAL ORIENTACION MANUFACTURA	4,759	84.42 %
INGENIERIA MECANICA	18,536	81.60 %
INGENIERIA MECANICA ORIENTACION MECATRONICA	2,923	82.70 %

Fig. 1 – Reporte desde DW representando Inscriptos a Cursada y Porcentaje de Regularización de Materias por carrera.

Esto permite, en una primera instancia, reconocer dos segmentos sujetos a análisis: el alumnado ingresante y el alumnado regular. A continuación, se muestra en un gráfico de barras la cantidad de Ingresantes a las distintas carreras en los últimos diez (10) años a la Facultad de Ingeniería.

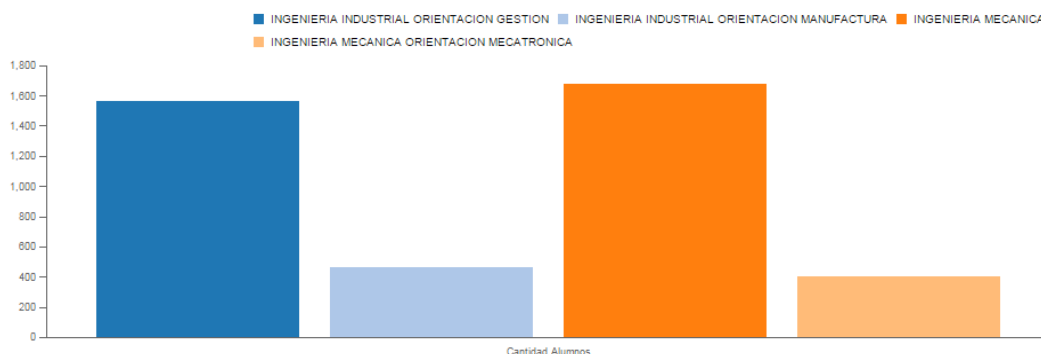


Fig. 2 – Gráfico de barras desde el DW de Cantidad de Ingresantes por Carrera a la Facultad de Ingeniería en los últimos diez (10) años.

A su vez, se evalúa el acceso al Campus Virtual por materia y la relación de rendimiento académico con el porcentaje de acceso a la plataforma.

Burch y Strater (1981) señalan que la información “debe tener tres características: ser oportuna, confiable y completa; para que pueda ser utilizada y genere ventajas competitivas”.

El primer eje competitivo de la información (oportuna) resulta sustentado mediante la aplicación tecnológica, ya que las autoridades de la institución pueden obtener la información que demanden de manera sencilla, en el momento que lo deseen mediante cualquier dispositivo desde la plataforma on-line.

El segundo eje competitivo (confiable) radica en el historial de datos de 10 años ininterrumpidos por lo que la información proporcionada es representativa de la situación actual de la institución y sujeta a estudio.

Finalmente, es en el tercer eje competitivo (completa) donde se insertan las nuevas dimensiones para consolidar la información del Data Warehouse necesario para la toma de decisiones de la Institución.

3 Conceptos generales

Para llevar adelante la implementación de un Data Warehouse (o cubo) es necesario definirlo. El mismo tiene un diseño como el que se muestra a continuación:

```
<Cube name="Finales" visible="true" cache="true" enabled="true">
  <Table name="rend_academico" schema="g2">
  </Table>
  <DimensionUsage source=
```

Se define el nombre del cubo, datos de conexión a la base de datos y dimensiones con sus respectivas métricas.

Se entiende por dimensión como la perspectiva de análisis y responden a la pregunta de cómo analizar el dato. La métrica representa el qué analizar (hechos) y se reconocen por ser cuantificables.

En este caso se define la dimensión de Carrera y la métrica que calcula cantidad de inscriptos, la cual realiza sumatoria de dicho dato.

```
<Dimension type="StandardDimension" visible="true" foreignKey="carrera"
highCardinality="false" name="Carrera">
```

```
  <Hierarchy name="Por carreras" visible="true" hasAll="true"
allMemberCaption="Todas las Carreras" allLevelName="Todas las carreras"
primaryKey="carrera">
```

```
    <Table name="sga_carreras" schema="g2">
```

```
  </Table>
```

```
    <Level name="Por Carreras" visible="true" table="sga_carreras"
column="nombre" type="String" uniqueMembers="false" levelType="Regular"
hideMemberIf="Never">
```

```
  </Level>
```

```
</Hierarchy>
```

```
</Dimension>
```

```
<Measure name="Inscriptos_Examen"
```

```
  column="inscriptos_examen" datatype="Integer" aggregator="sum"
caption="Inscriptos Examen" description="Cantidad de alumnos regulares que se
inscribieron a la Mesa de Examen." visible="true">
```

```
</Measure>
```

En los cubos se pueden definir métricas que son dinámicas y se basan en realizar operaciones matemáticas con los datos que se están visualizando y a partir de eso mostrar un resultado.

Por ejemplo, en uno de los cubos, el Porcentaje de Regularización de Cursada descontando a los alumnos ausentes se consigue según lo indica la fórmula (1):

$$\frac{\text{Aprobados de cursada}}{(\text{Inscriptos a la cursada} - \text{Ausentes a la cursada})} = \text{Porcentaje(\%)} \text{ de Regularización}$$

(1)

```
<CalculatedMember name="% Regularización sin Ausentes"
formatString="0.00 %" description="Se calcula el total de aprobados de la
cursada sobre el total de inscriptos a la cursada menos los ausentes"
formula="[Measures].[aprobados_cursada] / ([Measures].[Inscriptos_cursada] -
[Measures].[ausentes_cursada])" dimension="Measures" visible="true">
</CalculatedMember>
```

Otra de las métricas computadas es el Porcentaje de Desgranamiento de cursada, y se obtiene de la siguiente manera:

```
<CalculatedMember name="% Desgranamiento de Cursada" formatString="0.00 %"
description="Total de Alumnos que abandonaron sobre Total de Alumnos que se
inscribieron a la cursada" formula="[Measures].[ausentes_cursada] /
[Measures].[Inscriptos_Cursada]" dimension="Measures" visible="true">
</CalculatedMember>
```

4 Diseño e implementación

En las siguientes líneas, se pondrá en evidencia cómo se desarrollaron las nuevas dimensiones y métricas a fin de complementarse para proporcionar nueva información acerca del rendimiento académico de los alumnos. El objetivo de incorporar estos datos, tal como se explicó, resulta de la hipótesis de que tanto la forma en que los alumnos utilizan el Campus virtual y los docentes que están a cargo de las distintas cátedras, impactan y tienen incidencia directa sobre el rendimiento al cual se hace referencia.

4.1 Nuevas métricas y dimensiones

La primera métrica es la “Adaptabilidad del contenido curricular con el entorno virtual”. Esto implica una relación: “Campus Virtual – Materia”, mediante un tabla de ponderación que clasifica las materias de las carreras de grado analizadas según su posibilidad de ser tratadas a través del Campus Virtual. Es importante destacar que en general se visualiza una fuerte adaptabilidad al entorno virtual en

materias del ciclo superior como: Ingeniería Legal, frente a una baja adaptabilidad en materias del ciclo inicial como: Matemática I.

Las siguientes dos métricas son: “Utilización de la sección Foros” y “Utilización de la sección Correos”. Aquí, se establece una relación “Campus Virtual – Alumno”, mediante un promedio de intervenciones en foro y correos emitidos en una unidad de tiempo establecida.

Las últimas cuatro métricas son: “Utilización de la sección Presentación”, “Utilización de la sección Noticias”, “Utilización de la sección Archivos” y “Utilización de la sección Sitios”. En este caso, se evalúa la capacidad por parte de los responsables de la materia a utilizar todas las funcionalidades de la plataforma del Campus Virtual. Por lo tanto, la interpretación de esta métrica estará directamente asociada a la relación “Campus Virtual – Docente”.

La dimensión desarrollada es la denominada “Docentes” la misma consta de dos subdimensiones: “Nombre del docente” y “Cargo”. La misma tiene una función meramente descriptiva a fin de reconocer los docentes y su contexto en aquellas materias que se sometan a análisis por parte de las autoridades de la institución.

4.2 Nuevos reportes

En este apartado se ejemplifican algunos de los nuevos reportes que se pueden obtener a través del sistema DW.

En cuanto a gráficos, el sistema es capaz de generar gráficos de torta, de barras, de líneas y de tendencia.

A continuación, se muestra en un gráfico de torta la distribución de las materias según la “Adaptabilidad del contenido curricular con el entorno virtual”.

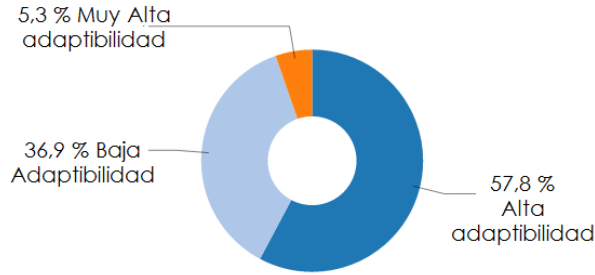


Fig. 3 – Gráfico de torta desde el DW de “Adaptabilidad del contenido curricular con el entorno virtual” de las materias de la Carrera de Ingeniería Industrial.

Para ejemplificar los gráficos de barra, se observa el grado de utilización de la sección Foros en materias escogidas arbitrariamente para ser sometidas a análisis.

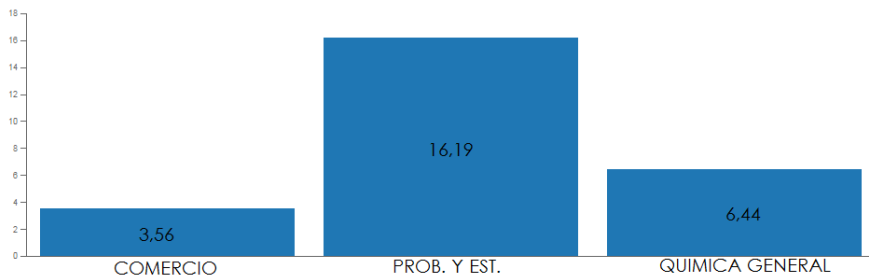


Fig. 4 – Gráfico de barra desde el DW de “Utilización de la sección Foros” de las materias Comercio, Probabilidad y Estadística y Química General.

Las primeras aproximaciones a la interpretación de la situación actual se basan en la intersección de los distintos datos aportados desde el DW.

En la figura 5, la tabla visualiza cómo en la dimensión “Docentes” se arroja la información de las personas a cargo de las diferentes cátedras, donde además, se incluye una tabla con los aprobados en las distintas materias.

Por Materias	Por Docente	Aprobados
ESTABILIDAD I	AREVALO	50
	CANEPA	50
	CAPUS	50
	DI BENEDETTO	50
	PENCO	50
MEDIOS DE REPRESENT. GRAFICA I	AMOROSO	52
	BLUMETTI	52
	BRUNO	52
	CASABURI	52
	DE PIERO	52
	FINGOLO	52
	GUIOTTO	52
	SAGGESE	52
	WYSOCKI	52

Fig. 5 – Reporte desde DW de los docentes a cargo de las cátedras de Estabilidad I y Medios de Representación Gráfica I junto con la cantidad de aprobados por cada materia.

Otra perspectiva de análisis se observa en la tabla de la figura 6, donde se reúnen los responsables de cada materia, la utilización de la sección Noticias y el porcentaje de aprobados de la misma.

Por Docente	Aprobados	Aula Virtual - Noticias (SI, NO)
AMOROSO	52	NO
AREVALO	50	SI
BLUMETTI	52	NO
BOFFI	63	SI
BRUNO	52	NO
CANEPA	50	SI
CAPUS	50	SI
CASABURI	52	NO
COZZA	63	SI
DASKEVICS	126	SI
DE PIERO	52	NO
DI BENEDETTO	50	SI
DUDELSAK	63	SI

Fig. 6 – Reporte desde DW de algunos docentes junto con el índice de utilización de la sección Noticias y la cantidad de aprobados por cada materia.

Conclusión

En materia de calidad educativa y procesos de enseñanza, es de común conocimiento que los índices de “gestión en la universidad” son los más sensibles cuando se someten a algún tratamiento. En su mayoría, esto se debe a la diversidad de factores que intervienen en dichos estudios: su cultura, las características de su población, las técnicas y tecnologías predominantes, etc. Es en este sentido que el trabajo presentado alcanzó los objetivos propuestos en la línea del mejoramiento del sistema de información para la toma de decisiones.

Con la incorporación de las nuevas métricas y dimensiones se aumentaron las fuentes que nutren los datos actuales, de esta manera se completó y consolidó el DW, posicionándolo como una herramienta fehaciente para que las autoridades analicen la situación real de la Institución.

Por otro lado, es inevitable observar como la realización de este tipo de estudios impacta rápidamente en los docentes, generando la intención de autoevaluaciones sobre el trabajo en sus cátedras.

Como líneas futuras de trabajo se propone el análisis de los resultados obtenidos en cada Carrera debido a que los índices entregados por el software permiten que las materias sean estudiadas dentro de una población de materias comparables según su tipología. Por consiguiente, estudiar el redimiendo académico de los alumnos de esta facultad abarca no solo el historial académico de los alumnos si no también la realidad actual de los docentes (como agentes formadores) y de la tecnología (como instrumento de formación).

Alineados con la Misión de la Facultad de Ingeniería de UNLZ, el objetivo de los futuros trabajos continúan en la línea de alcanzar excelencia en el nivel educativo del proceso enseñanza/aprendizaje.

Bibliografía

William H. Inmon: Building the Data Warehouse. John Wiley & Sons, 4th Edition, New York (2005) - ISBN: 978-0-7645-9944-6

Thomas H. Davenport - Lawrence Prusak: Working Knowledge: How Organizations Manage What They Know. Harvard Business School Press, Boston (1998)

El portal educativo del Estado argentino - Educación virtual, <http://portal.educ.ar/>

Fernandez, B., Suarez, L., ALVÁREZ, E.: El camino hacia el Espacio Europeo de Educación Superior: deficiencias metodológicas y propuestas de mejora desde la perspectiva del alumno. *Aula Abierta*. Vol88, 85—105 (2006)

Izquierdo, J.M., Pardo, M.L.: Las Tecnologías de la Información y las Comunicaciones (TIC) en la gestión académica del proceso docente educativo en la educación superior. *Revista Pedagógica Universitaria*. Vol. XII. Nº 1, 58--68 (2007)

Hamidian, B.: Usos y necesidades de formación en tecnología de información y comunicación de los docentes de la Facultad de Ciencias Económicas y Sociales (Universidad de Carabobo). Tesis doctoral inédita, Universidad de Sevilla (2010)

Quijano, S.: *Organizaciones: ¿Qué son? ¿Cómo han sido pensadas? Su comportamiento. Su evaluación.* Editorial Bárdenas, Barcelona (2000)

Burch, J.G., Strater, F.R.: *Information Systems: Theory and practice.* John Wiley & Sons, New York (1981)