

Aplicando estrategias y tecnologías de Inteligencia de Negocio en sistemas de gestión académica

Francisco Javier Díaz, M. Alejandra Osorio, A. Paola Amadeo, Dalila L. Romero
Laboratorio de Investigación en Nuevas Tecnologías Informáticas.
Facultad de Informática. Universidad Nacional de La Plata
Calle 50 esq. 120, 2do Piso
+54 221 4223528
jdiaz@unlp.edu.ar, {aosorio, pamadeo, dalila_r}@cespi.unlp.edu.ar

Resumen

El siguiente artículo presenta una línea de investigación que tomo identidad propia a partir del año 2011, como resultado de actividades previas relacionadas con la aplicación de herramientas y técnicas de inteligencia de negocio a datos almacenados en los sistemas académicos de gestión universitaria, que se utilizan para la operativa diaria en las unidades académicas. Todos ellos basados en software libre o con acceso. El cruce con la información relevada de la actividad de los estudiantes en las redes sociales abre un campo de investigación muy interesante, para analizarla en conjunto con su desempeño académico, su comportamiento en las Plataformas Virtuales de Aprendizaje, en el uso de la biblioteca, etc.

Este análisis permitirá analizar el perfil del estudiante de la Facultad de Informática de la UNLP, carreras y planes de estudio implementados, la actividad de los docentes en relación al desenvolvimiento de las materias que dictan, permitiendo realizar acciones para resolver distintas situaciones identificadas a partir de datos confiables y fiables. Esta línea de investigación ha promovido la propuesta de dictado de una materia optativa relacionada con el tema a partir del año 2011, con amplia aceptación por parte de los estudiantes y una tesina de Licenciatura.

Palabras clave: Inteligencia de Negocios, Sistemas de Gestión Académica, Redes Sociales

Contexto

El Laboratorio de Investigación en nuevas Tecnologías Informáticas de la Facultad de Informática de la Universidad Nacional de La Plata incluye dentro de sus líneas de investigación el área de Análisis de la Información. Esta área intenta dar respuesta a la creciente demanda de analistas de información que permitan generar conocimiento a partir de los sistemas de gestión operativa que se utilizan hoy en día en toda organización, con distintos fines.

En particular, en la Facultad de Informática se utiliza el sistema SIU Guaraní desde el año 2005, gestionando la información académica de más de 4.151 alumnos en el año 2011, más de 700 ingresantes y graduados según el Anuario Estadístico de la UNLP [1] El proyecto de Data Warehouse del Ministerio de Educación de la Nación [2] provee un sistema de análisis de información de sistemas académicos, basado inicialmente en O3 [3] y en Pentaho BI [4], una solución open source para inteligencia de negocios.

Asimismo se utiliza el sistema de código abierto Moodle [5] para la gestión de cursos virtuales que complementan la actividad presencial. Este sistema está integrado con el sistema SIU Guaraní a través de una interfaz común, que facilita la gestión de las inscripciones a trabajos prácticos entre ambos sistemas. Este desarrollo fue propuesto inicialmente por el Consorcio SIU del Ministerio de Educación de la Nación [6], y la integración con Moodle 2 fue desarrollada por el equipo de implementación de Sistemas

Académicos del Centro Superior para el Procesamiento de la Información [7] de la UNLP. El sistema Moodle se utiliza en la Facultad desde el año 2005, incluyendo más de 310 cursos y más de 14000 usuarios. Puede consultarse a través de <http://catedras.info.unlp.edu.ar>, <http://cursos.linti.unlp.edu.ar> y <http://postgrado.linti.unlp.edu.ar>. La información almacenada en este sistema involucra entregas de tareas, participación en los foros, acceso a los diferentes recursos y demás registros los cuales, al cruzarlos con el desenvolvimiento académico pueden resultar en datos significativos para la toma de decisiones de los directivos de la Facultad.

Por otra parte, para la gestión de la biblioteca se utiliza el sistema de software libre Meran [8], desarrollado por el equipo de desarrollo del CeSPI a partir de Koha [9], el sistema de SL para la gestión de bibliotecas implementado en distintas facultades de la UNLP a partir del año 2003. Este sistema permite gestionar los procesos bibliotecarios y los servicios a los usuarios, como estantes virtuales para las cátedras, la posibilidad de votar un libro o dar recomendaciones. Todo esto integrado también al sistema SIU Guaraní.

Finalmente, la Facultad de Informática de la UNLP está haciendo uso también de las redes sociales para poder llegar a los distintos sectores de la sociedad que están vinculados con ella y de esta manera realizar una mejor difusión de las actividades que se llevan a cabo. En particular dispone de un perfil en Facebook,

<http://www.facebook.com/facultad.informatic.a.unlp> y distintas cuentas de difusión académica como @InformaticaUNLP con más de 1100 seguidores, @infounlp con más de 500 seguidores entre otras. Además ciertas cátedras utilizan estos canales para comunicarse con sus estudiantes, como la cátedra de Algoritmos y Estructuras de Datos @ayed_fi, Introducción a los Sistemas Operativos @iso_info_unlp con más de 100 seguidores, Sistemas Operativos, Tecnologías Aplicadas a Business Intelligence

@catedraBI con más de 20 seguidores, entre otros.

Como se puede observar, se cuentan con distintas fuentes de información relacionadas con el mismo sujeto, el estudiante, que consolidadas en único repositorio permitirá realizar análisis transversales, que sirvan de insumo a grupos interdisciplinarios con distintos perfiles.

Introducción

Business Intelligence, es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios. Asociándolo directamente con las tecnologías de la información, podemos definir Business Intelligence como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la organización) en información estructurada, para su explotación directa (reporting, análisis OLTP / OLAP, alertas...) o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre el negocio[10]. Por su parte, un Data Warehouse se constituye como una parte fundamental de BI que surge a partir de la necesidad de las organizaciones de contar con información homogénea y confiable, en forma global e integrada. Facilitan la construcción y mantenimiento de datos estructurados que permiten realizar análisis de los datos en un entorno diferente al operacional. Este análisis promueve la transformación de los datos en información y la información en conocimiento.[11]

Asimismo, la tendencia en el avance de la tecnología que ha abierto las puertas hacia un nuevo enfoque de entendimiento y toma de decisiones, la cual es utilizada para describir enormes cantidades de datos (estructurados, no estructurados y semi estructurados) que tomaría demasiado tiempo y sería muy costoso cargarlos a un base de datos relacional para su análisis. De tal manera que,

el concepto de Big Data aplica para toda aquella información que no puede ser procesada o analizada utilizando procesos o herramientas tradicionales. Esta tecnología se encuentra en pleno desarrollo, encontrando soluciones open source que es necesario estudiar e investigar en forma sistemática para obtener resultados comparativos que sean de utilidad. Podemos mencionar aquí herramientas que realizan analytics en memoria, por ejemplo el streaming processing que realizan empresas como LinkedIn, Groupon a través de aplicaciones como Storm[12] y Kafk[13]; o Drill[14] y Dramel[15] para la exploración de datos. D3[16] es otra aplicación muy poderosa para crear tableros interactivos en forma rápida y eficaz visualmente.

Analizar y aplicar técnicas y estrategias, evaluar distintas herramientas de software libre disponibles en el mercado, relacionadas con BI y Big Data es la propuesta principal de este trabajo.

Asimismo, el estudio de las mismas requieren conocimientos previos de bases de datos básico. Es interesante entonces trabajarlos también con estudiantes de grados avanzados de la carrera, que hacen uso de buscadores de internet, de libros en línea, de las redes sociales como herramienta motivador y que permita reflejar emociones, compartir y aplicar técnicas del aprendizaje en red, reflexionando asimismo sobre su propio aprendizaje. El conectivismo como teoría de aprendizaje[17][18] y la aplicación de técnicas de learnign analytics[19] se constituyen como últimas tendencias en educación[20], sobre las que es necesario reflexionar y hacer experiencias que reflejen la forma de aprender de nuestros estudiantes.

Líneas de Investigación y Desarrollo

La tarea de investigación que estamos llevando a cabo comprenden varias líneas de trabajo todas ellas en el área de Inteligencia de Negocios y en el aprendizaje de estas temáticas.

Por un lado, aplicar distintas metodologías para la constitución de un Data Warehouse que soporte las distintas estrategias y técnicas para obtener conocimiento a partir de los datos almacenados en distintos sistemas. Que contemple la información de todas las fuentes de información mencionadas previamente. La construcción de un DW involucra una serie de actividades relacionadas con distintas técnicas para la Extracción, Transformación y Carga soportadas por distintas herramientas como Kettle de Pentaho BI, Spago BI [21], entre otras. La naturaleza del negocio determina los requerimientos de los usuarios que en este caso están relacionados con rendimiento académico, comportamiento en entornos virtuales y redes sociales.

Respecto al primer punto, rendimiento académico, se está trabajando en los cubos de análisis de información provistos por el Ministerio de Educación de la Nación, así como también en el diseño e implementación de un cubo de desgramiento por cohorte que permitan abordar la problemática de deserción en la Facultad y se puedan definir distintas políticas o estrategias para trabajarla.

Durante el año 2011 y 2012 se trabajó con la Facultad de Ingeniería de la UNLP, intentando dar respuesta a los requerimientos referidos al análisis del rendimiento de determinadas materias, cohortes, docentes y años académicos. El trabajo realizado permitió a la Facultad realizar un análisis pormenorizado de sus alumnos a lo largo de la carrera y trabajar en conjunto con el MEN para lanzar el Plan Estratégico de Ingeniería 2012 - 2016 [22]

El trabajo conjunto con la Facultad de Ciencias Económicas ha permitido implementar programas como de caracterización de aspirantes y del seguimiento del graduado, para el diagnóstico y comprensión de las características de los alumnos de esta Facultad. [23]

Por otro lado, la aplicación de técnicas de minería de datos permitirá identificar los diferentes perfiles de los estudiantes y egresados, ayudando a comprender mejor su comportamiento en las plataformas virtuales y las redes sociales, para implementar

propuestas educativas que los asistan para completar sus estudios, con el objetivo de disminuir la deserción en las carreras de Informática [24]. El grado de aprovechamiento de la información minada por parte del usuario final depende en gran medida de una correcta visualización y una interfaz amigable de interacción por este motivo se prevé el desarrollo de aplicaciones ad hoc para facilitar las consultas de los usuarios finales. Distintas herramientas de minería brindan librerías que se integran con entornos de desarrollo para simplificar esta tarea.

La aplicación de técnicas y el análisis y uso de herramientas de Big Data se constituye como otra línea de investigación por sí misma. Durante el año 2012 en la materia Tecnologías Aplicadas para Business Intelligence se introdujo el tema y los alumnos investigaron sobre las distintas herramientas, profundizando sobre una de ellas. El trabajo realizado se puede consultar en <http://www.facebook.com/catedrabi.unlp>. Poder extender el análisis a datos de toda la universidad y facilitar la integración las bases de datos abiertas de organismos públicos permitirá aplicar las técnicas y herramientas analizadas sobre esta temática. Analizar el comportamiento de los estudiantes en las redes sociales constituirá otro caso de estudio.

El uso de redes sociales para la enseñanza de TI con alumnos avanzados permite realizar primeras experiencias en la aplicación de nuevos paradigmas educativos, como el conectivismo y el aprendizaje en red [25].

Resultados y Objetivos

Entre los objetivos planteados podemos mencionar:

- Brindar herramientas que favorezcan la toma de decisiones gerenciales trabajando para la mejora continua que se vean reflejados en cambios permanentes para la institución.

- Transformar toda la información contenida en los sistemas operacionales, sobre los cuales se tiene acceso, de una manera rápida, consistente, íntegra y confiable, colaborando con ayudar a los directivos para propiciar mejoras en las ofertas educativas y los servicios ofrecidos por la institución.
- Analizar herramientas para aplicar técnicas de data analytics y Big data sobre las redes sociales, como Netvizz [26], Google Analytics y Gephi [27]
- Innovar en el dictado de materias de años avanzados de la carrera acorde a los nuevos paradigmas educativos y formas de comunicarse de la sociedad.

Entre los resultados obtenidos al momento podemos mencionar que se ha afianzado el dictado de la materia optativa “Tecnologías Aplicadas a BI” dado que la matrícula va en aumento y se incorporaron dos colaboradores en este último año. Asimismo la tasa de asistencia es alta lo cual es significativo en alumnos que son activos laboralmente en un 80%.

Se han realizado convenios de capacitación sobre la temática con organismos públicos provinciales, como la Dirección Provincial de Sistemas de Información y Tecnología de la provincia de Buenos Aires.

Se trabaja en conjunto con el Ministerio de Educación de la Nación en el desarrollo de nuevos cursos e implementación sobre sistemas de SL.

Formación de Recursos Humanos

La capacitación permanente en temas relacionados con la temática ha permitido afianzar el grupo de trabajo de 2 profesores, 1 ayudante y 2 colaboradores, que en conjunto aprenden y colaboran en la experimentación y

construcción de casos de estudio. El trabajo conjunto con el personal del Centro Superior para el Procesamiento facilita y promueve la implementación de las técnicas y estrategias propuestas.

Se encuentran en desarrollo una tesina de Licenciatura en Informática denominada Aplicación de técnicas y estrategias de Inteligencia de Negocio (BI) para analizar/integrar información de los alumnos de la Facultad de Informática de la UNLP. Y se trabaja con distintos alumnos interesados en realizar su tesina sobre la temática.

Referencias

- [1] Anuario Estadístico UNLP 2012: Indicadores Comparados. <http://www.unlp.edu.ar/indicadores> (Ultima visita Marzo 2013)
- [2] <http://www.siu.edu.ar/>
- [3] <http://www.ideasoft.biz/>
- [4] <http://www.pentaho.com/>
- [5] <http://www.moodle.org/>
- [6] Consorcio SIU <http://.siu.edu.ar> (Ultima visita Marzo 2013)
- [7] <http://www.cespi.unlp.edu.ar/>
- [8] Meran UNLP <http://www.meran.unlp.edu.ar/> (Ultima visita Marzo 2013)
- [9] <http://www.koha.org/>
- [10] <http://www.bi-argentina.com.ar/que-es-business-intelligence/>
- [11] Kimball 1992
- [12] storm-project.net
- [13] <http://kafka.apache.org/design.html>
<http://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/index.html>
- [14] <http://gigaom.com/cloud/for-fast-interactive-hadoop-queries-drill-may-be-the-answer/>
- [15] http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en/us/pubs/archive/36632.pdf
- [16] <http://d3js.org>
- [17] <http://www.elearnspace.org/Articles/connectivism.htm>
- [18] <http://aprenderapensar.net/2012/12/10/el-conectivismo-como-forma-de-aprendizaje/>
- [19] <http://www.solaresearch.org/>
- [20] <http://blogcued.blogspot.com.ar/2011/10/es-el-conectivismo-una-teoria-lo-es-del.html>
- [21] www.spagobi.org
- [22] http://www.ing.unlp.edu.ar/institucional/difusion/2012/ingreso_retencion
- [23] http://www.econo.unlp.edu.ar/caracterizacion_aspirantes
- [24] Tesina de Licenciatura en Informática de A.C. Kruzylko Ostojic, Claudia Yanina en desarrollo
- [25] www.fernandosantamaria.com
- [26] <https://lab.digitalmethods.net/~brieder/facebook/netvizz/>
[fb_source=bookmark_apps&ref=bookmarks&count=0&fb_bmpos=3_0&state=cddffb9cc8ddcabfa791ca4a2df113f5&code=AQB5eo6CLldv_Dj0BSAUwApJDFhwSQBw1kwIrnoZD0pNqxwoy0lu05Ys8uPDqw8GefvHtGADhDeaKHYAGC9bnb8HTEeabHW6wbWNqIsFzLMhUARwN3ZFwd6yOgOGZURPM3PN5Qjz-T9oFnuADbE9IHVwVRzDUvIMK620GY4qFX9pgdutky0qRjqibaMj7-rGX59w8VVk2wbaIY79xpuEHUG#_=_](https://www.facebook.com/#!/page.php?fb_source=bookmark_apps&ref=bookmarks&count=0&fb_bmpos=3_0&state=cddffb9cc8ddcabfa791ca4a2df113f5&code=AQB5eo6CLldv_Dj0BSAUwApJDFhwSQBw1kwIrnoZD0pNqxwoy0lu05Ys8uPDqw8GefvHtGADhDeaKHYAGC9bnb8HTEeabHW6wbWNqIsFzLMhUARwN3ZFwd6yOgOGZURPM3PN5Qjz-T9oFnuADbE9IHVwVRzDUvIMK620GY4qFX9pgdutky0qRjqibaMj7-rGX59w8VVk2wbaIY79xpuEHUG#_=_)
- [27] <http://www.whatsnew.com/2011/09/29/gephi-impresionante-plataforma-de-visualizacion-de-datos-de-codigo-abierto/>