

Ciencia de datos en bibliotecas populares sanjuaninas

Prog. Luis Olguin, Mag. Alejandra Malberti, Mag. Raúl Klenzi

Instituto de Informática – Departamento de Informática

Facultad de Ciencias Exactas, Físicas y Naturales Universidad Nacional de San Juan

lolguin@iinfo.unsj.edu.ar; amalberti@gmail.com; rauloscarklenzi@gmail.com

RESUMEN

Las *Bibliotecas Populares* son creadas, gestionadas y mantenidas por los propios asociados. Son entidades autónomas (ONGs) que adaptan sus acciones a las necesidades de su público. Existen en Argentina desde hace más de 150 años, y han ido adecuándose a la situación social, política y económica del país. Son, sin lugar a dudas, un lugar de "acceso abierto" al conocimiento, asumiendo el papel de mediador entre el saber y el ciudadano común.

Actualmente las bibliotecas populares de San Juan, que poseen sistemas de gestión automatizada, no elaboran estrategias de acción que fortalezcan su accionar **analizando los datos** que el sistema de gestión genera, perdiendo una oportunidad de descubrir particularidades de sus visitantes que les permitan mejorar su posición en el contexto social.

La presente propuesta, encarada a través del proyecto *Ciencia de Datos en Bibliotecas Populares Sanjuaninas (CiDaBi-SJ)*, pretende abordar el *paradigma de la Ciencia de Datos* centrándose en el análisis, extracción de conocimiento y valor asociado al conjunto de datos generados por los distintos sistemas de gestión bibliotecaria que poseen las bibliotecas populares de San Juan.

Se indaga en cómo adaptar la *analítica de datos* a las necesidades de una biblioteca

popular, estudiando cada caso y aportando soluciones realistas.

Palabras clave: Ciencia de Datos, Biblioteca Popular, Visualización de Datos

CONTEXTO

Esta investigación se enmarca dentro de la *convocatoria de Vinculación Tecnológica de la Secretaría de Políticas Universitarias "Agregando Valor 2017" (RS-2017-31806560-APN-SECPU#ME)* y forma parte de las líneas abordadas por el *Laboratorio de Sistemas Inteligentes para la Búsqueda de Conocimiento en Datos Masivos* y el *Laboratorio de Bibliotecas Digitales* del Instituto de Informática de la FCEFN (*Resol. 2/2015-CD-FCEFN*). Los docentes que integran el equipo de CiDaBi-SJ vienen trabajando en temáticas relacionadas desde el año 2005. Como antecedentes se encuentran los proyectos “*Descubrimiento de conocimiento a través de Data Warehousing y Data Mining, en los datos de la Biblioteca de la Facultad de Ciencias Exactas, Físicas y Naturales*”, 21/E639, desarrollado en el período 2005-2007; “*Búsqueda estratégica de conocimiento en los datos de biblioteca y de alumnos de la FCEFN*”, 21/E824, período 2008-2010; “*Minería de datos en la determinación de patrones de uso y perfiles de usuarios*”, 21/E889, período 2011-2013, “*Extracción de Conocimiento en Datos*

Masivos, periodo 2014-2015 y "Visualización y Deep Learning en Ciencia de Datos", periodo 2018-2019.

1. INTRODUCCIÓN

El trabajo que llevan adelante las bibliotecas populares sanjuaninas, que poseen sistemas de gestión automatizados incluye la aplicación de *técnicas de análisis de la información* que estos sistemas generan - en general por falta de formación de RRHH en estas temáticas- encontrándonos con *bibliotecarios operadores del sistema*, es decir adaptados a lo que los diferentes módulos ofrecen (catalogación, circulación, asociación, etc.) sin tener como objetivo la *búsqueda de patrones ocultos y útiles o resultados estadísticos novedosos*.

La acción de usar un sistema de gestión bibliotecaria, no implica la extracción de patrones automáticamente y por tanto se desperdicia la oportunidad de usar esas *pepitas de información* para la mejora de otros servicios como ser la *diseminación selectiva de la información* entre los asociados a una biblioteca popular (Ruiz-Lobaina, Esther Marina y Pedro Lázaro Romero-Sánchez. 2017).

En la actualidad *parece natural* que las bibliotecas exploren el uso de las herramientas de ciencia de datos para satisfacer la demanda creciente de sus usuarios, tal es el caso de la biblioteca de Signapure que haciendo uso de herramientas de *big data* ha logrado optimizar los procesos de selección y oferta de obras a sus usuarios, tarea que antes se efectuaba solo a *criterio humano*. Esto se logra a través del *análisis de demanda*, técnica que permite pronosticar la

demanda de los usuarios, aumentando en consecuencia la satisfacción del asociado (SINGAPORE LIBRARY MINES BIG DATA, 2014).

El análisis de grandes datos sin dudas es un paso que los profesionales de las bibliotecas deben dar para poder extraer información valiosa de sus repositorios (Alonso Arévalo, Julio; Vázquez Vázquez, Marta. 2016).

En San Juan lamentablemente no se aborda, desde el Instituto de Formación en Bibliotecología dependiente de Ministerio de Educación de la provincia, la formación de RRHH en estas técnicas, por lo que el abordaje de la temática a través de la experiencia del equipo de trabajo de IdeI sin duda será un aporte para las bibliotecas populares involucradas.

En la actualidad el enunciado "los datos masivos están en el punto de remodelar nuestro modo de vivir, trabajar y pensar" (Mayer-Schönberger, 2013) debe ser analizado seriamente por las bibliotecas y, dado que el conocimiento (o parte de él) se almacena en ellas, contar con herramientas que permitan extraer *información de calidad, oportuna y adecuada* es una necesidad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La posibilidad de almacenar enormes volúmenes de datos exige el estudio de nuevas formas de representarlos para extraer información valiosa. En el proyecto CiDaBi-SJ se propone el estudio de herramientas y técnicas que permitan analizar grandes volúmenes de datos y generar representaciones visuales que posibiliten la toma de decisiones tanto a los bibliotecarios

como a los usuarios finales de una biblioteca popular.

La línea principal de investigación, que se ajusta a una investigación en informática aplicada, consiste en el empleo de soluciones, especialmente software libre, para la toma de decisiones en el ámbito de una biblioteca popular.

3. RESULTADOS OBTENIDOS / ESPERADOS

El proyecto Ciencia de Datos en Bibliotecas Populares Sanjuaninas (CiDaBi-SJ) es un proyecto de investigación aplicada que se sustenta en la hipótesis que “las bibliotecas populares deben asumir la aplicación de la Ciencia de Datos para obtener una mejora en la calidad de sus servicios y procesos que les permita mejorar su competitividad en un marco de sostenibilidad y actuación socialmente responsable”.

Para llevar a cabo la hipótesis se precisará definir indicadores y variables que ayuden a establecer si es factible la aplicación de las técnicas de ciencia de datos en una biblioteca popular.

Se espera que al finalizar este proyecto sea posible mostrar patrones de uso de la biblioteca popular por parte de los usuarios que a ella acuden desde el enfoque de la ciencia de datos, y que estos resultados sean interpretados de manera correcta tanto por parte de los bibliotecarios como del público en general, es decir los destinatarios de esta investigación.

La ejecución del proyecto CiDaBi-SJ permitirá a los alumnos participantes adiestrarse en los procesos de extracción de datos independientemente de su fuente, la

limpieza de los datos eliminando los atributos que no aportan valor a los resultados buscados, dominar técnicas de procesamiento de datos usando diferentes métodos y algoritmos de minería de datos, y la visualización gráfica (cualitativa y cuantitativamente) de los datos y resultados.

A los fines de alcanzar los objetivos propuestos para CiDaBi-SJ se prevé como eje principal la participación activa de la entidad adoptante, tomando en cuenta los medios y recursos disponibles. Se buscará gestionar de manera sistemática el compromiso de la entidad adoptante para encontrar un balance entre sus expectativas y los objetivos del proyecto. Los estudiantes se verán favorecidos por el nuevo conocimiento además de la posibilidad concreta de transferencia al medio. Las temáticas abordadas en el proyecto brindarán a los docentes participantes nuevas líneas de investigación a abordar, así como la posibilidad de realizar aportes al conocimiento en las áreas relacionadas.

4. FORMACIÓN DE RRHH

Olguín, Luis Alberto. Integrante del proyecto. Es actualmente maestrando de la Maestría en Informática de la Universidad Nacional de San Juan, siendo el título de su trabajo de tesis Red de Co-Préstamo en Bibliotecas

Formación de alumnos avanzados de las carreras Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información, en el uso de herramientas de la ciencia de datos.

Formación de docentes de las carreras del Departamento de Informática, integrantes del

proyecto, en el tratamiento de datos relativos a la problemática abordada.

Entrenamiento a los RRHH de la entidad adoptante en el uso de herramientas informáticas, por medio de las capacitaciones previstas en el plan de trabajo de CiDaBi-SJ.

content/uploads/2014/09/BI_Asean_Sept_2014_FINAL.pdf> [Consulta: 08-02-2019].

5. REFERENCIAS

Alonso Arévalo, Julio; Vázquez Vázquez, Marta (2016). *Big Data: la próxima «gran cosa» en la gestión de la información* - . BiD: textos universitarios de biblioteconomía i documentació , núm. 36 (jun) . <<http://bid.ub.edu/es/36/alonso.htm>>. [Consulta: 10-02-2019].

Mayer-Schönberger, (2013). *Big Data: La revolución de los datos masivos*. Madrid: Turner. <<http://catedradatos.com.ar/media/3.-Big-data.-La-revolucion-de-los-datos-masivos-Noema-Spanish-Edition-Viktor-Mayer-Schonberger-Kenneth-Cukier.pdf>>. [Consulta: 10-02-2019].

Ozdemir, S. (2016). *Principles of Data Science*. Packt Publishing Ltd.

Ruiz-Lobaina, Esther Marina y Pedro Lázaro Romero-Sánchez, (2017). *Búsqueda de patrones para mejorar productos y servicios en las bibliotecas*. Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información 72 (31): 209-225. <<http://dx.doi.org/10.22201/iibi.0187358xp.2017.72.57830>>.[Consulta: 10-02-2019].

Singapore Library Mines Big Data (2014). Business Information, ASEAN, vol. 2. <<http://www.asiandatasience.com/wp->