

Cómo citar el artículo

Marulanda, C., López, M. & Mejía, M. (2017). El acompañamiento educativo como estrategia de cercanía impulsadora del aprendizaje del estudiante. *Revista Virtual Universidad Católica del Norte*, 50, 224-237. Recuperado de <http://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/view/821/1339>

Minería de datos en gestión del conocimiento de pymes de Colombia*

Carlos Eduardo Marulanda Echeverry

Ingeniero Industrial
Especialista en diseño y manufactura asistida por computador
Magíster en administración
Candidato a doctor en ingeniería, industria y organizaciones
Profesor del Departamento de Sistemas e Informática de la Universidad de Caldas y del Departamento de Administración de la Universidad Nacional de Colombia, Sede Manizales
carlose@ucaldas.edu.co, cemarulandae@unal.edu.co

Marcelo López Trujillo

Ingeniero de Sistemas
Magíster en educación
Doctor en ingeniería informática, sociedad de la información y del conocimiento
Profesor del Departamento de Sistemas e Informática de la Universidad de Caldas y del Departamento de Informática y Computación de la Universidad Nacional de Colombia, Sede Manizales
mlopez@ucaldas.edu.co y mlopeztr@unal.edu.co

María Helena Mejía Salazar

Ingeniera de Sistemas
Magister en Ciencias Computacionales
Doctora en ciencias computacionales
Profesora del departamento de sistemas e informática de la Universidad de Caldas y del departamento de informática y computación de la Universidad Nacional de Colombia, Sede Manizales.
mariah.mejia@ucaldas.edu.co y mhmejiasa@unal.edu.co

Recibido: 22 de julio de 2016.

Evaluado: 22 de julio de 2016.

Aprobado: 27 de marzo de 2017.

Tipo de artículo: investigación científica y tecnológica.

* Investigación: proyecto de investigación titulado *Aplicación de técnicas de minería de datos a estudios del estado de la gestión del conocimiento en pymes de Colombia*, código: 1746214. Grupo GITIR, Departamento de Ingeniería, Universidad de Caldas. Calle 65 n.º 26-10, Manizales (Colombia).



Resumen

La importancia de la minería de datos para descubrir patrones y tendencias desde la gestión del conocimiento (GC) en las empresas de hoy, es fundamental dadas las situaciones de competitividad y desarrollo que deben enfrentar éstas. Se presenta en éste artículo la aplicación de técnicas de minería de datos a una base de datos de GC de 322 PYMES de las ciudades de Bogotá, Medellín, Cali, Barranquilla, Bucaramanga, Manizales, Pereira y Armenia, para lo cual se utilizaron enfoques investigativos como el exploratorio y el descriptivo, conjuntamente con técnicas de minería de datos y el uso de una herramienta de software propietaria. De lo cual se destaca el avance en competencias, herramientas y TIC para la GC.

Palabras clave

Egresados, Minería de datos, técnicas de clasificación

Data mining in knowledge management of SMEs in Colombia

Abstract

The Importance of Data Mining to uncover patterns and trends from Knowledge Management (KM) in today's companies is fundamental given the situations of competitiveness and development that these must face. This article presents the application of data mining techniques to a KM database of 322 SMEs in the cities Bogotá, Medellín, Cali, Barranquilla, Bucaramanga, Manizales, Pereira and

Armenia. For which we used research approaches such as exploratory and descriptive organized with data mining techniques and the use of a proprietary software tool.

As result, some aspects like the progress in skills, tools and ICT for the KM were highlighted.

Keywords

Graduates, Data mining, Classification techniques.

L'exploration de données en gestion des connaissances des PME en Colombie

Résumé

L'importance de l'exploration de données pour découvrir les modèles et les tendances de la gestion des connaissances (GC) dans les entreprises aujourd'hui est essentiel étant donné des situations de compétitivité et de développement qu'elles affrontent. Dans cet article se présente l'application des techniques d'exploration de données GC à une base de données de 322 PME aux villes Bogotá, Medellín, Cali, Barranquilla, Bucaramanga, Manizales, Pereira et Armenia, pour lesquels des approches de recherche telles qu'exploratoire et descriptive ont été utilisés conjointement avec des techniques d'exploration de données et l'utilisation de l'outil logiciel propriétaire. Des quels se souligne l'avance dans les compétences, les outils et les TIC pour la GC.

Mots Clés

Les diplômés, l'exploration de données, les techniques de classification.

Introducción

El proceso dinámico de la gestión del conocimiento (GC) usualmente empieza con la creación, descubrimiento y recolección interna de conocimiento y de las mejores prácticas, seguido por compartir y entender las prácticas que la organización puede usar, y de esta manera ajustar y aplicar dichas prácticas a nuevas situaciones, mejorando el desempeño organizacional (O'Dell & Grayson, 2004, citado por Zhao, 2010).

Al respecto, Alavy y Leidner (2001) consideran algunos procesos básicos de creación, almacenamiento y recuperación, transferencia y aplicación del conocimiento.

- Creación del conocimiento: implica el desarrollo de nuevos contenidos, o remplazar el contenido existente dentro del conocimiento tácito y explícito de la organización a través de procesos de colaboración, así como procesos cognitivos individuales.
- Almacenamiento y recuperación: también conocidos como 'memoria de la organización', constituyen un aspecto importante del conocimiento de la organización y su efectiva gestión.
- Transferencia y aplicación de conocimiento, que se produce en diversos niveles: la transferencia de conocimientos entre los individuos, de los individuos a fuentes explícitas, de personas a los grupos, entre los grupos y del grupo de la organización. La aplicación se da en términos de cinco elementos: 1), valor percibido de la fuente de la unidad conocimiento; 2) motivacional, para compartir el conocimiento; 3), la existencia y la riqueza de los canales de transmisión; 4), motivacional para recibir; y 5), la capacidad de absorción de la unidad receptora, que se define como la capacidad no solo para adquirir y asimilar, sino también para utilizar el conocimiento.

En este sentido, Natek y Zwilling (2014) explican que en el proceso de GC, la técnica de minería de datos (MD) puede ser utilizada para extraer y descubrir conocimiento valioso y significativo de una gran cantidad de datos. Hoy, la MD se ha reconocido como una herramienta de análisis emergente y parte de la estructura de la GC, asociada con la inteligencia empresarial, la cual trae una nueva sinergia y valor añadido a los datos. Algunas de las técnicas de MD utilizadas en el medio se relacionan con analizar factores de influencia clave, detectar categorías, pronosticar, resaltar excepciones, analizar escenarios y hacer predicciones de cálculo, entre otras labores.

Natek y Zwilling (2014) continúan explicando los pasos para el análisis de la MD, así: el primer paso es la creación de los datos conjuntos; el segundo, definir la herramienta de MD a utilizar; en tercer lugar, evaluar las técnicas de MD a utilizar; y, por último, analizar los datos por cada modelo y elegir el mejor.

Tsai (2013) complementa lo anterior explicando que la MD es un campo interdisciplinario que combina inteligencia artificial, gestión de bases de datos, visualización de datos, aprendizaje automático y algoritmos matemáticos y estadísticos. Esta tecnología ofrece diferentes metodologías para la toma de decisiones, resolución de problemas, el análisis, la planificación, el diagnóstico, la detección, la integración, la prevención, el aprendizaje y la innovación.

Con lo anterior, MD se define como un proceso que utiliza matemática, inteligencia artificial estadística y técnicas de aprendizaje automático para extraer e identificar información útil. La mayoría de las técnicas de GC y MD implican patrones

de aprendizaje, edificados sobre el fundamento de datos o información existente y de aprendizaje automático, y la inteligencia artificial.

Gestión del conocimiento

La GC se presenta como una disciplina cuyo objetivo se centra en desarrollar el conocimiento en las fases de adquisición, almacenamiento, transformación, distribución y utilización, con la finalidad de lograr ventajas competitivas (Riesco, 2006; Barney, 1991; Dosi, 1992). Los usos de la GC y las razones que los motivan son variados: van desde generar cambios y resultados sustentables, hasta optimizar recursos, aprovechar el conocimiento existente, aprender permanentemente y estimular la creatividad e innovación (Rodríguez, 2006), así como aumentar las oportunidades de nuevos productos y servicios; mejorar la comunicación interna y con los grupos de interés; mejorar la competitividad; resaltar el liderazgo de las organizaciones en su sector; y mejorar los niveles de productividad y rendimiento, aprovechando las competencias de los empleados.

Desde un enfoque fundamentado en los recursos internos de la organización como base para obtener una ventaja competitiva, Barney (1991) establece que la información y el conocimiento se han convertido en factores claves de las organizaciones exitosas. Barnes (2002), a su turno, define que, desde el planteamiento basado en el conocimiento, el conocimiento organizativo se convierte en el capital más rentable.

Lo anterior se ratifica con lo establecido por autores como Arthur y Huntley (2005), Collins y Smith (2006), Mesmer-Magnus y DeChurch (2009), y Lin (2007), que en sus trabajos investigativos muestran que la combinación e intercambio de conocimientos en una organización, se relaciona directamente con la disminución de costos; la creatividad e innovación de productos; el mejoramiento organizacional; y el aumento de su rendimiento y de los ingresos por ventas (Barragán & Zubieta, 2006).

Según Tsoukas (2001), el conocimiento es la capacidad del individuo para establecer distinciones dentro de un dominio de acción, con base en una apreciación de contexto o teoría —o de ambos—; y el conocimiento organizacional es la capacidad que los miembros de una organización se han desarrollado para hacer distinciones en el proceso de llevar a cabo su trabajo en contextos concretos.

La GC es el proceso que continuamente asegura el desarrollo y aplicación de todo tipo de conocimientos pertinentes en una empresa, con el objeto de mejorar su capacidad de resolución de problemas y así contribuir a la sostenibilidad de sus ventajas competitivas (Andreu & Siever, 1999). Alavy y Leidner (2001) definen la GC en organizaciones como aquella que se refiere a la identificación y el aprovechamiento del conocimiento colectivo en una organización, en aras de ayudarla a competir.

Gold, Malhotra y Segars (2001) proponen que los elementos claves para determinar el éxito o fracaso de las iniciativas de gestión del conocimiento se encuentran en la identificación y valoración de los prerrequisitos que son necesarios para que dicho esfuerzo prospere. Dichos prerrequisitos son descritos con frecuencia como “recursos” y “capacidades” en la literatura del comportamiento organizacional (Kelly & Amburgey, 1991; Law, Wong & Mobley, 1998; Leonard, 1995). Earl (2001) complementa lo anterior considerando que la GC puede ser coherente con la teoría de recursos y capacidades, es decir, construir y competir en una capacidad que podría ser bastante difícil de imitar.

La GC se define como “el proceso de gestión de la información como un recurso estratégico para mejorar el desempeño organizacional” (Chaffey & Wood, 2005; citado por Mikroyannidis & Theodoulidis, 2010) (Alavi & Leidner, 1999; citado por Kankanhalli & Tan, 2005).

La GC apunta a proveer la información correcta a la persona correcta en el momento correcto (Bassi, 1999; Broadbent, 1998; citados por Zhao, 2010); su objetivo es hacer que la organización tenga mayor poder competitivo, elasticidad e innovación a través de los miembros de la organización. Se trata de un concepto combinado basado en los sistemas basados en conocimiento, inteligencia artificial, mejoras en la ingeniería del *software*, gestión de los recursos humanos y comportamiento organizacional (Liebowitz, 2000; citado por Zhao, 2010). No solo se trata de recordar y compartir conocimiento, sino también de marginar, descartar y olvidar el conocimiento que no haya sido calificado como relevante o legítimo (Gherardi & Nicolini, 2002; citados por Ferguson, Huysman & Soekijad, 2010).

228

Minería de datos

Se puede definir la Minería de Datos (MD) como el proceso de descubrir conocimiento útil y entendible, desde grandes bases de datos almacenados en distintos formatos, por medio de modelos inteligibles a partir de los datos.

Según Pérez-Palacios et al. (2014), la MD es una parte importante de un proceso más amplio conocido como descubrimiento de conocimiento en bases de datos (KDD en inglés). El objetivo principal de la MD consiste en la extracción de información oculta de un conjunto de datos. Esto puede ser alcanzado por el análisis automático o semiautomático de gran cantidad de datos, lo que permite la extracción de patrones desconocidos. Estos pueden ser grupos de registros de datos (análisis clúster), inusuales registros (detección de anomalías) y las dependencias entre datos (asociación reglas). Por lo tanto, los patrones pueden ser vistos como un resumen de los datos de entrada, y se pueden utilizar para su posterior análisis.

Ahora bien, según Caridad (2008), el concepto de MD representa una idea que ha venido madurando durante muchos años, en el sentido de recorrer grandes bases de datos para recuperar información conceptual de interés e inferir nueva

información útil. La MD es un proceso posterior, destinado a lograr un mejor conocimiento de la información disponible, aumentar beneficios o ventas, y disminuir pérdidas; es decir, tiene un objetivo distinto al que ha motivado la recogida y almacenamiento de información.

Liao, Chen y Deng (2010) concluyen que la MD es el proceso de descubrimiento significativo del conocimiento del cliente —patrones, asociaciones, cambios y estructuras de grandes cantidades de datos almacenados en bases de datos—, desde un enfoque que integra la minería matemática, de datos tradicional, y técnicas evolutivas con un sistema multiagente.

De acuerdo con Peña-Ayala (2014), la MD es un proceso dedicado a escanear enormes repositorios de datos para generar información y descubrir conocimiento. La MD persigue descubrir patrones de datos, organizar la información, definir reglas y estructuras de asociación, estimar elementos desconocidos, clasificar objetos, y desvelar muchos tipos de resultados que no se producen fácilmente; de este modo, los resultados de la MD representan un valioso apoyo para la toma de decisiones.

En esencia, existen dos tipos de modelos de MD: descriptivos y predictivos. Los primeros suelen aplicar funciones de aprendizaje sin supervisión para producir patrones que explican o generalizar la estructura intrínseca, las relaciones y la interconexión de los datos extraídos. Los segundos, por su parte, se aplican con frecuencia al aprendizaje supervisado para calcular valores desconocidos o futuros de las variables dependientes, con base en las características de las variables independientes relacionadas.

Muhammad, Mohamudally y Babajee (2013) explican que el foco de la extensa investigación en el campo de la MD está en el desarrollo y mejora de los algoritmos existentes, así como en la evaluación de los conocimientos descubiertos como un proceso de un solo paso o de múltiples pasos, desde diferentes enfoques —como el álgebra relacional y la teoría de la información, entre otros—.

Riquelme, Ruiz y Gilbert (2006), al igual que Liao, Chu y Hsiao (2012) definen las tareas más comunes de la minería de datos:

Clasificación: clasifica un dato dentro de una de las clases categóricas predefinidas. (...)

Regresión: el propósito de este modelo es hacer corresponder un dato con un valor real de una variable. (...)

Clustering: se refiere a la agrupación de registros, observaciones, o casos en clases de objetos similares. Un clúster es una colección de registros que son similares entre sí, y distintos a los registros de otro clúster. (...)

Generación de reglas: aquí se extraen o generan reglas de los datos. Estas reglas hacen referencia al descubrimiento de relaciones de asociación y dependencias funcionales entre los diferentes atributos. (...)

Resumen o sumarización: estos modelos proporcionan una descripción compacta de un subconjunto de datos. (...)

Análisis de secuencias: se modelan patrones secuenciales, como análisis de series temporales, secuencias de genes, etc. El objetivo es modelar los estados del proceso, o extraer e informar de la desviación y tendencias en el tiempo. (...) (Riquelme, Ruiz & Gilbert, 2006, p. 13).

Señalan, igualmente, que el rápido crecimiento del interés en la minería de datos se debe

(i) al avance de la tecnología de Internet y a la gran participación en aplicaciones multimedia en este dominio, (ii) a la facilidad en la captura de datos y el abaratamiento de su almacenaje, (iii) a compartir y distribuir los datos en la red, junto con el aumento de nuevas bases de datos en los repositorios, (iv) al desarrollo de algoritmos de aprendizaje automático robustos y eficientes para procesar estos datos, (v) al avance de las arquitecturas de las computadoras y la caída del coste del poder computacional, permitiendo utilizar métodos computacionalmente intensivos para el análisis de datos, (vi) la falta de adaptación de los métodos de análisis y consulta convencionales a nuevas formas de interacción y finalmente (vii) a la potencia que este tipo de análisis vienen mostrando como herramientas de soporte a la toma de decisiones frente a realidades complejas (...). (Riquelme, Ruiz & Gilbert, 2006, p. 14).

230

A la par con lo anterior, Ur-Rahman y Harding (2012) explican que la tecnología de MD proporciona flexibilidad para explotar la información desde múltiples formatos o bases de datos, tales como bases de datos relacionales, almacenamiento de datos y bases de datos transaccionales, etc. Esto es complementado por PhridviRaj y GuruRao (2014), en tanto comentan que la diferencia entre los datos en las bases de datos y un almacén de datos es que, en las primeras, los datos están en la forma estructurada; mientras que en segundo pueden o no estar presentes de tal manera. La estructura de los datos puede ser definida para que sea compatible para su procesamiento; de ahí que en la MD se debe iniciar con un proceso de la limpieza de los datos, con el fin de hacer factible su posterior procesamiento. Los pasos para el desarrollo de la MD incluyen la preparación de los datos, el modelamiento, la evaluación, el desarrollo y el mejoramiento de medición.

Descripción de la base de datos

La base de datos está estructurada con respuestas a preguntas en escala Likert, calificadas de 1 a 5: 1 corresponde a "se está en desacuerdo" o "no realizado"; 2,

“realizado parcialmente”; 3, “realizado en intervalos”; 4, “realizado con regularidad”; y 5, “realizado completamente”. Se agruparon las preguntas en las dimensiones y categorías explicitadas en la tabla 1.

Tabla 1. Estructura de la base de datos

Dimensión	Categorías	Número de preguntas
Análisis Organizacional	Análisis de conocimiento, análisis de debilidades, amenazas, fortalezas y oportunidades, DAFO, establecimiento de objetivos, selección de la estrategia, definición de factores clave e indicadores.	27
Competencias	Gestión de la información y la documentación, gestión de la comunicación, diseño de herramientas digitales, gestión de la innovación y el cambio, y gestión del aprendizaje organizacional.	44
Prácticas	Identificar, generar, retener, compartir y aplicar conocimiento.	33
Procesos	Consideraciones generales, procesos para identificar, generar, retener, compartir y aplicar conocimiento.	52
TIC	Web social y semántica para identificar, generar, retener, compartir y aplicar conocimiento; TIC para el modelo SECI (socializar, exteriorizar, combinar e interiorizar conocimiento) de Nonaka y Takeuchi (1999); TIC para un modelo holístico de gestión de conocimiento.	61

Fuente: elaboración propia.

Método

Contando con la base de datos mencionada, y con la necesidad de extraer información y conocimiento considerando los datos almacenados en ella, se hizo uso de una aplicación de *software* propietaria que permite utilizar las tareas más corrientes en la MD (Riquelme, Ruiz & Gilbert, 2006; Liao, Chu & Hsiao, 2012), a saber:

- Etapa de selección: Se obtuvieron los repositorios de datos. Como fuente interna se seleccionó la base de datos de la GC de las pymes de Colombia. Como fuente externa se seleccionaron el cuestionario completo con las dimensiones, variables y categorías.
- Etapa de procesamiento y limpieza: se procedió a eliminar los valores indeterminados o nulos, y se estandarizaron valores que estaban incoherentes y afectaban el patrón de calidad de la base de datos.
- Etapa de transformación de datos: se discretizaron algunos de los atributos con valores continuos y se transformaron los valores numéricos en valores discretos o nominales, con el fin de disminuir el número de valores distintos de estos atributos.

- Etapa de minería de datos: para el efecto se utilizó una herramienta de *software* propietaria que permitió desarrollar la clasificación. Igualmente, se desarrolló un modelo entidad – relación que se explicita en la figura 1.

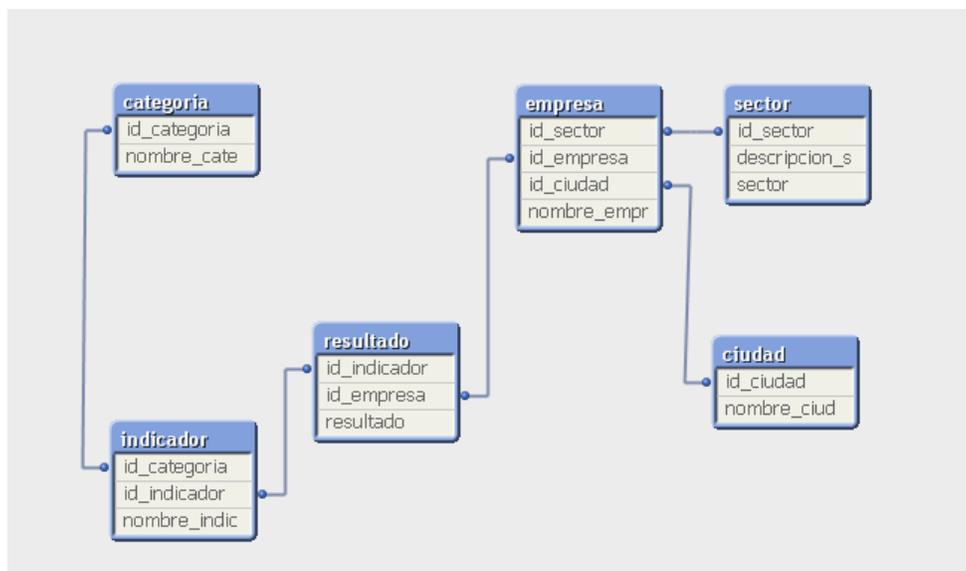


Figura 1. Modelo entidad – relación. Fuente: elaboración propia.

El número de empresas por ciudad que fueron consideradas se muestra en la figura 2.

Ciudades de estudio	Total Empresas Evaluadas	Empresas Armenia	Empresas Barranquilla	Empresas Bogota	Empresas Bucaramanga
9	321	29	27	52	25
	Empresas Cali	Empresas Dos Quebradas	Empresas Medellin	Empresas Manizales	Empresas Pereira
	15	3	38	107	26

Figura 2. Total de empresas. Fuente: elaboración propia.

Resultados y discusión

Luego de seguir las etapas para el desarrollo de la MD y diseñar el modelo entidad – relación, se procedió a usar la aplicación propietaria de *software* sobre la base de datos de GC de pymes de Colombia, con lo que se encontraron los resultados que se exponen en esta sección.

En lo relacionado con las buenas prácticas para la GC que realizan las pymes de las ciudades evaluadas, se obtuvieron los resultados que se observan en la figura 3.

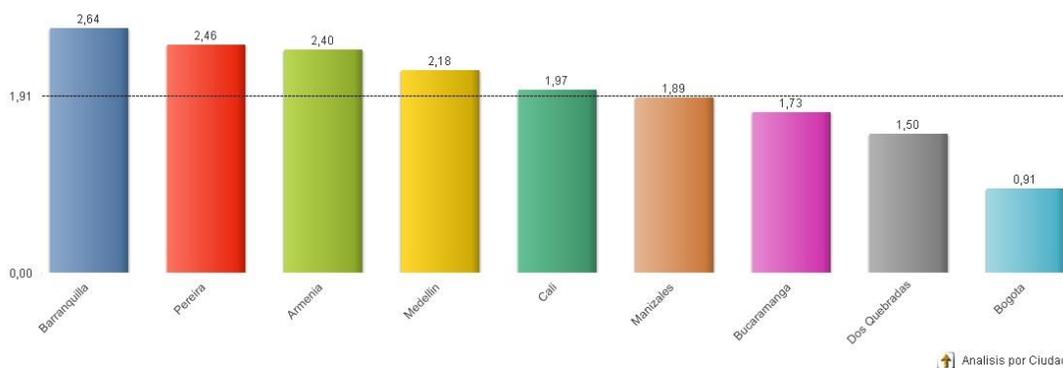


Figura 3. Buenas prácticas de GC por ciudad. Fuente: elaboración propia.

De acuerdo a la calificación, se observa que se está desarrollando una cantidad más elevada de buenas prácticas en Barranquilla y muy pocas en Bogotá. Asimismo, se destacan a este respecto ciudades como Pereira y Armenia, mientras que se está haciendo poco en Dosquebradas y Bucaramanga. 4 ciudades están por debajo del promedio y 5 por encima del promedio de calificación. Las prácticas de mayor auge son las relacionadas con la generación, retención y el compartir el conocimiento

En lo relacionado con las competencias personales para desarrollar la GC que contemplan las pymes de las ciudades evaluadas, se obtuvieron los resultados que se observan en la figura 4.

233

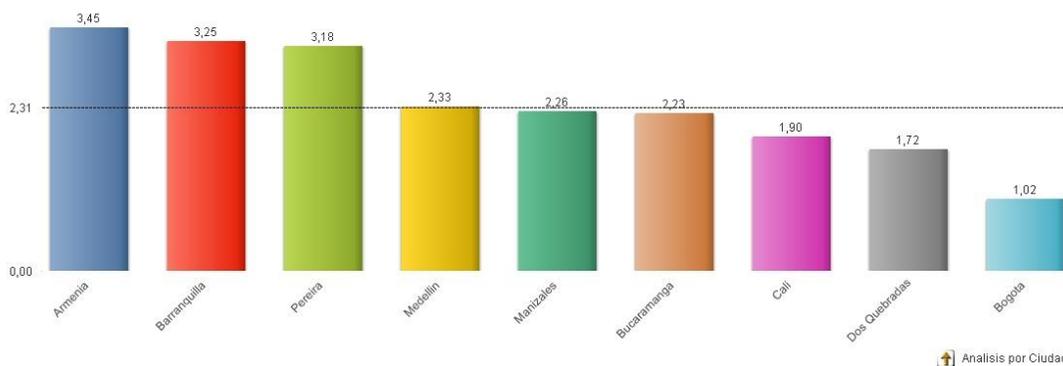


Figura 4. Competencias para la GC por ciudad. Fuente: elaboración propia.

De acuerdo a la calificación, se están generando más competencias para la GC en la ciudad de Armenia, y pocas en Bogotá; se destacan ciudades también Barranquilla y Pereira, mientras se está generando poco en Dosquebradas y Cali. 5 ciudades están por debajo del promedio y 4 por encima del promedio de calificación. Las competencias más utilizadas son las relacionadas con el aprendizaje organizacional y la gestión de la información y la comunicación, así como la efectiva documentación de la GC.

En lo relacionado con procesos para la GC que contemplan las pymes de las ciudades evaluadas, se obtuvieron los resultados que se observan en la figura 5.

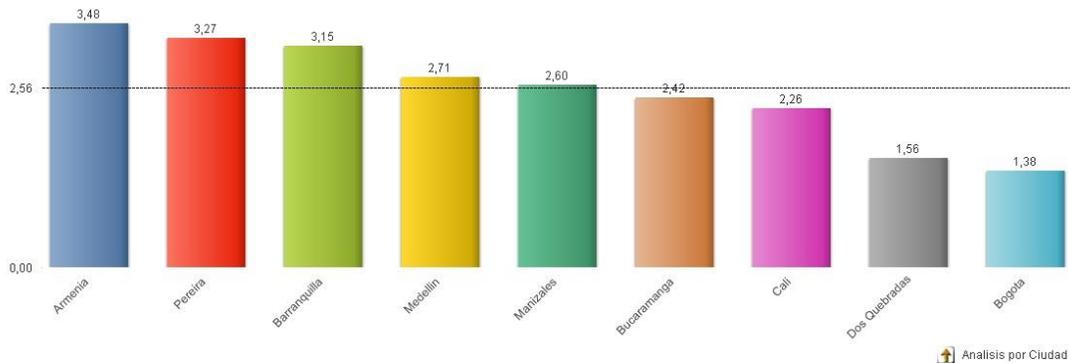


Figura 5. Procesos para la GC por ciudad. Fuente: elaboración propia.

Del mismo modo que en el punto anterior, se están generando más procesos para la GC en Armenia, mientras que en Bogotá son escasos. Igualmente, Pereira y Barranquilla se destacan en esta materia, mientras que se generan pocos procesos en Dosquebradas y Cali. 4 ciudades están por debajo del promedio y 5 por encima del promedio de calificación. Los procesos más utilizados para la GC son los relacionados con la identificación y la retención del conocimiento.

234

Conclusiones

La técnica de MD implica, en sus primeras fases, limpieza, ajuste y transformación, de acuerdo a las necesidades del estudio. Estos procesos son necesarios porque, a pesar de todos los controles para la gestión de la información, los usuarios incurren de modo permanente en errores de digitación que hacen difícil la coincidencia de los datos y la futura gestión del conocimiento.

Las pymes evaluadas deben avanzar mucho más en el uso de buenas prácticas para la GC, así como en el desarrollo de más y mejores competencias personales y organizacionales para la GC, sin dejar de lado el uso y apropiación de procesos mucho más refinados, incluso desde el aporte de las TIC para una GC efectiva.

Referencias

- Alavi, M. & Leidner, D. E. (1999). Knowledge Management Systems: Issues, Challenges and Benefits. *Communications of AIS*, 1(2es), 1-37.
- Alavy, M. & Leidner, D. (2001). Review knowledge management and knowledge management systems conceptual foundations and research issues. *MIS Quarterly*, 107-136.

- Andreu, R. & Siever, S. (1999). La gestión integral del conocimiento y del aprendizaje. *Economía Industrial*, 68, 63-72.
- Arthur, J. B. & Huntley, C. L. (2005). Ramping up the organizational learning curve: Assessing the impact of deliberate learning on organizational performance under gainsharing. *Academy of Management Journal*, 1159–1170.
- Barnes, S. (2002). *Sistemas de Gestión del Conocimiento, teoría y práctica*. Argentina: Thomson.
- Barney, J. (1991). Firm Resources and sustained competitive advantage. *Journal de of management*, 99-100.
- Barragan, O. A., & Zubieta, G. J. (2006). *La administración de la calidad como generador de innovación dentro de las organizaciones*. Madrid: Memorias del I Congreso Iberoamericano de Ciencia, Tecnología, Sociedad e innovación.
- Bassi, L. J. (1999). Harnessing the Power of Intellectual Capital. *Training and Development*, 51(12), 25-30.
- Broadbent, M. (1998). The Phenomenon of Knowledge Management: What does it Mean to the Information Profession. *Information Outlook*, 2(5), 23-36.
- Caridad, J. O. (2008). LA MINERÍA DE DATOS: ANÁLISIS DE BASES DE DATOS EN LA EMPRESA.
- Chaffey, D., & Wood, S. (2005). *Business Information Management: Improving Performance Using Information Systems*. FT Prentice Hall.
- Collins, C. J., & Smith, K. G. (2006). Knowledge exchange and combination: The role of human resource practices in the performance of high-technology firms. *Academy of Management Journal*, 544–560.
- Dosi, G. T. (1992). Toward a theory of corporate coherence: preliminary remarks. *Technology and Enterprise in a Historical Perspective*.
- Earl, M. (2001). Knowledge Management Strategies: Toward a Taxonomy. *Journal of Management information Systems*, 215 - 233.
- Ferguson, J., Huysman, M. & Soekijad, M. (2010). Knowledge Management in Practice: Pitfalls and Potentials for Development. *World Development*, 38(12), 1797-1810.
- Gherardi, S., & Nicolini, D. (2002). Learning in a Constellation of Interconnected Practices: Canon or Dissonance? *Journal of Management Studies*, 39(4), 419-436.
- Gold, A. H., Malhotra, A., & Segars, A. H. (2001). Knowledge Management: An Organizational Capabilities Perspective. *Journal of Management Information Systems*, 18(1), 185-214.
- Kankanhalli, A., & Tan, B. C. (2005). Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation. *MIS Quarterly*, 29(1), 113-143.
- Kelly, D., & Amburgey, T. (1991). Organizational Inertia and Momentum: A Dynamic Model of Strategic Change. *Academy of Management Journal*, 34(3), 383-397.
- Law, K. S., Wong, C. & Mobley, W. H. (1998). Toward a Taxonomy of Multidimensional Constructs. *Academy of Management Review*, 23(4), 741-753.

- Leonard, D. (1995). *Wellsprings of Knowledge: Building and Sustaining the Source of Innovation*. Boston: Harvard Business School Press.
- Liao, S.-h., Chen, Y.-J. & Deng, M. (2010). Mining customer knowledge for tourism new product development and customer relationship management. *Expert Systems with Applications*(37), 4212–4223.
- Liao, S.-H., Chu, P.-H. & Hsiao, P.-Y. (2012). Data mining techniques and applications – A decade review from 2000 to 2011. *Expert Systems with Applications*(39), 11303–11311.
- Liebowitz, J. (2000). *Building Organizational Intelligence: A Knowledge Management Primer*. London: CRC Press.
- Lin, H. F. (2007). Knowledge sharing and firm innovation capability: An empirical study. . *International Journal of Manpower*, 315–332.
- Mesmer-Magnus, J. R. & DeChurch, L. A. (2009). Information sharing and team performance: A meta-analysis. . *Journal of Applied Psychology*, 535–546.
- Mikroyannidis, A., & Theodoulidis, B. (2010). Ontology Management and Evolution for Business Intelligence. *Internationa Journal of Information Management*, 30, 559-566.
- Muhammad, D. K., Mohamudally, N., & Babajee, D. (2013). A Unified Theoretical Framework for Data Mining. *Procedia Computer Science*(17), 104 – 113.
- Natek, S. & Zwilling, M. (2014). Student data mining solution–knowledge management system related to higher education institutions. *Expert Systems with Applications*(41), 6400–6407.
- Nonaka, I. & Takeuchi, H. (1999). *La organización creadora de conocimiento*. México D.F.: Oxford University Press.
- O'Dell, C. & Grayson, C. J. (2004). Identifying and Transferring Internal Best Practices. En *Handbook on Knowledge Management* (págs. 601-622). Spring Berlin Heidelberg.
- Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications*(41), 1432–1462.
- Pérez-Palacios, T., Caballero, D., Caro, A., Rodríguez, P. & Antequera, T. (2014). Applying data mining and Computer Vision Techniques to MRI to estimate quality traits in Iberian hams. *Journal of Food Engineering*(131), 82–88.
- PhridviRaj, M. & GuruRao, C. (2014). Data mining – past, present and future – a typical survey on data streams. *Procedia Technology*(12), 255 – 263.
- Riesco, M. (2006). *El negocio es el conocimiento*. Madrid: Díaz de Santos.
- Riquelme, J., Ruiz, R. & Gilbert, K. (2006). Minería de Datos: Conceptos y Tendencias. *Revista Iberoamericana de Inteligencia Artificial*, 11-18.
- Rodríguez G. (2006). Modelos para la creación y gestión del conocimiento: una aproximación teórica. *Educar*, 25-39.
- Tsai, H.-H. (2013). Knowledge management vs. data mining: Research trend, forecast and citation approach. *Expert Systems with Applications*(40), 160–3173.

- Tsoukas, H. (2001). What is organizational Knowledge. *Journal of management studies*, 973-993.
- Ur-Rahman, N., & Harding, J. (2012). Textual data mining for industrial knowledge management and text classification: A business oriented approach. *Expert Systems with Applications*(39), 4729–4739.
- Zhao, J. (2010). School Knowledge Management Framework and Strategies: The New Perspective on Teacher Professional Development. *Computers in Human Behavior*, 26(2), 168-175.