

# Uso de herramientas de minería de datos y de inteligencia tecnológica para identificar patrones de publicación y patentamiento en entornos nacionales, así como sus concordancias

Área de investigación: Administración de la Tecnología

**Roberto E. López-Martínez**  
Instituto de Ingeniería  
Universidad Nacional Autónoma de México  
México  
[robertol@unam.mx](mailto:robertol@unam.mx)

## CONGRESO INTERNACIONAL DE CONTADURÍA ADMINISTRACIÓN E INFORMÁTICA



Octubre 5, 6 y 7 de 2011  
Ciudad Universitaria  
México, D.F.

<http://congreso.investiga.fca.unam.mx>

[informacongreso@fca.unam.mx](mailto:informacongreso@fca.unam.mx)

Teléfonos

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax

52 (55) 5616.03.08



**ANFECA**  
Asociación Nacional de Facultades y  
Escuelas de Contaduría y Administración

División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

# Uso de herramientas de minería de datos y de inteligencia tecnológica para identificar patrones de publicación y patentamiento en entornos nacionales, así como sus concordancias

## Resumen

El presente trabajo, de corte experimental, busca la aplicación de algunas herramientas de inteligencia tecnológica-competitiva y de minería de datos para la identificación de patrones que revelen tanto la estructura de la investigación científica como de la investigación aplicada, así como su concordancia en el entorno de un país. Este tipo de información puede ser de gran utilidad para la toma de decisiones de políticas de fomento a la ciencia la tecnología y la innovación, en distintos niveles de decisión (universitaria, empresarial, regional y nacional). Se parte de una hipótesis experimental referida a la existencia de concordancia entre la estructura de investigación científica y la de patentamiento a nivel de país; es decir, en un país determinado hay una similitud entre la organización de lo que se investiga y lo que se patenta. Para llevar a cabo el trabajo experimental se emplean datos recuperados de la base de datos del *Institute for Scientific Information* y de la plataforma *Espacenet* de la Oficina Europea de Patentes, para estudiar el caso de México. Posteriormente los datos son procesados mediante diversas técnicas de minería de datos y de análisis semántico, así como procedimientos estadísticos. Se concluye que mediante el uso de este tipo de métodos es posible identificar el alineamiento o la falta de éste entre las temáticas de investigación con las problemáticas de interés industrial.

**Palabras clave:** *inteligencia tecnológica, minería de datos, política de ciencia y tecnología*

Octubre 5, 6 y 7 de 2011  
Ciudad Universitaria  
México, D.F.

<http://congreso.investiga.fca.unam.mx>

[informacongreso@fca.unam.mx](mailto:informacongreso@fca.unam.mx)

Teléfonos

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax

52 (55) 5616.03.08



División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

## INTRODUCCIÓN

Dentro de las esferas de teoría y aplicación relativas a las políticas de ciencia, tecnología e innovación (CTeI) y particularmente en lo que respecta al análisis de los sistemas de innovación, puede resultar importante conocer el grado de alineamiento entre las actividades de investigación científica y las de aplicación tecnológica industrial. Esto es, dentro del análisis y evaluación sobre la conformación de un sistema de innovación es crucial conocer si existe una correspondencia entre los esfuerzos de investigación y los directamente asociados con la generación de innovaciones. Una forma de aproximarse a este problema es a través de la recuperación y análisis de publicaciones científicas y documentos de patente generados en un país o una región.<sup>1</sup> Para ello se emplean diversas técnicas de recuperación de información y de métrica científica y tecnológica que son también comunes a las actividades de inteligencia tecnológica competitiva (ITC).

El propósito de este trabajo es experimentar con métodos, procedentes de áreas de informática relacionadas con la minería de datos y la búsqueda y recuperación de información — específicamente con el análisis semántico latente (ASL), para complementar las técnicas de ITC y así obtener una herramienta más poderosa para identificar la concordancia entre las actividades de investigación y las de patentamiento. Existen casi nulos antecedentes de trabajo en esta área, salvo un trabajo reciente que explora con fines distintos, la confiabilidad del ASL para detectar similitudes entre documentos de patentes y publicaciones científicas (Magerman et al., 2010), aunque empleando muestras de textos más reducidas de las que presentamos en este trabajo. A continuación describiremos brevemente los principios conceptuales de la ITC, particularmente del uso de técnicas de co-ocurrencia de términos así como la idea básica y procedimiento del análisis semántico latente. Posteriormente dedicamos una sección a la metodología desarrollada para el estudio exploratorio, para pasar finalmente a la descripción de resultados y a las conclusiones del trabajo.

Octubre 5, 6 y 7 de 2011

### Inteligencia tecnológica competitiva

En términos generales, el “ciclo de inteligencia” puede ser definido como el proceso mediante el cual colecciones de datos son recuperadas, reunidas, transmitidas, evaluadas, analizadas y transformadas en información útil para la toma de decisiones y la acción. Ahora bien, de manera más específica, la inteligencia competitiva se refiere a información operativa acerca del entorno de negocios, que podría tener algún impacto en la posición competitiva de una empresa, es decir, información que va más allá de los meros descubrimientos sobre el ambiente de negocios e incluye recomendaciones concretas para dar respuesta a las observaciones, análisis y conclusiones (Ashton & Klavans, 1997). Un tipo particular de inteligencia competitiva se dirige hacia la ciencia y tecnología (C&T) como componentes de los negocios, constituyendo el campo de la ITC que enfatiza por un lado, la función de investigación y desarrollo de una organización y por el otro abarca también a otras actividades asociadas al desarrollo tecnológico, tales como la planeación tecnológica estratégica y los procesos de adquisición de tecnología y de inversión en equipamiento entre otros.

Las actividades de ITC usualmente se dirigen a proporcionar una alerta temprana sobre desarrollos externos que representen amenazas y oportunidades potenciales; a la evaluación

<http://informacongreso@fca.unam.mx>

Teléfonos<sup>1</sup>

Véase por ejemplo la contribución de este autor en (UNCTAD, 2011, capítulo 2 y anexos).

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax

52 (55) 5616.03.08

División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

ANFECA  
Asociación Nacional de Facultades y  
Escuelas de Contaduría y Administración

oportuna de innovaciones y prospectos de colaboración generados por actividades externas de manera que sea posible preparar las respuestas apropiadas; y, a la anticipación y comprensión de los cambios y tendencias de la ciencia y la tecnología en el entorno competitivo, como preparación para la planeación organizacional y el desarrollo de estrategias. Estos objetivos son alcanzados mediante una vigilancia y análisis sistemáticos que incluyen diversos esfuerzos de recolección y filtrado de datos, tales como el “barrido” (estudio amplio del entorno), el monitoreo (seguimiento rutinario y enfocado hacia tópicos científicos y tecnológicos de interés específico) y la exploración (colección y filtrado de datos sobre tecnologías, expertos u organizaciones particulares).

### ***Métrica científica y análisis de co-ocurrencia de términos***

Varios métodos que originalmente se desarrollaron con el propósito de evaluar la investigación a través del empleo de las bases de datos de publicaciones que comenzaron a generarse desde mediados de la década de los cincuentas, han sido paulatinamente incorporados a la ITC. Nos enfocaremos ahora, en el análisis de co-ocurrencia de términos y las técnicas de mapeo asociadas al mismo, como medios para revelar la estructura cognitiva de la investigación.

Durante los últimos sesenta años diversas teorías sobre la construcción del conocimiento científico y la innovación han sido desarrolladas tanto desde algunas corrientes de pensamiento económico como desde la sociología de la ciencia y la tecnología. Varias de éstas incluyen o están basadas en la idea de las redes socio-cognitivas y socio-técnicas y han conducido a una serie de herramientas matemáticas y modelos para explicar la producción de conocimiento y consecuentemente, de tecnología.<sup>2</sup> Desde esta perspectiva, la innovación puede ser estudiada a través de la red de alianzas que la hacen posible, lo que implica que toda la ciencia y la tecnología están unidas por una dinámica de redes. Así, la fusión de los enfoques socio-cognitivo y técnico, asume que es posible identificar la red de un laboratorio o una empresa a través de sus publicaciones y registros de patentes y así obtener una visión comprehensiva de su futura producción de este tipo de documentos, o al menos tener cierta idea sobre el posible contenido temático de éstos. De manera similar, cuando publicaciones o patentes de un campo específico son colectadas y analizadas, es posible identificar la red que constituye la estructura cognitiva del mismo obteniéndose una cierta visión prospectiva sobre el conocimiento que será desarrollado dentro del mismo.

El interés de explotar los datos contenidos en bases tales como la del *Institute for Scientific Information* (ISI) para la elaboración de mapas de la ciencia, data del final de la década de los setentas; esta idea implicaba el tratar a las áreas de conocimiento como objetos o entidades con una ubicación específica en un espacio de  $n$  dimensiones en el cual la distancia entre objetos es significativa y bien definida. Por tanto, la ‘cartografía científica’ comenzó como un intento de lograr una representación física de campos y disciplinas —y en un nivel menor de agregación, de artículos individuales, para los cuales la ubicación relativa de las entidades es mostrada (Garfield et al., 1978). Desde esta propuesta, las relaciones entre los campos científicos y tecnológicos, así como su evolución temporal han interesado a diversas organizaciones. Los métodos y técnicas derivados para obtener representaciones cartográficas de estas relaciones han sido usados para propósitos diversos, incluyendo la mercadotecnia, la administración, la planeación, la mejora de

<http://congreso.investiga.fca.unam.mx>

[informacongreso@fca.unam.mx](mailto:informacongreso@fca.unam.mx)

Teléfono<sup>2</sup> Ver por ejemplo (Callon et al., 1986) y (Molina, 1990).

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax 52 (55) 5616.03.08

División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

la comunicación entre los agentes involucrados en el desarrollo de actividades científicas y tecnológicas y en general para la inteligencia tecnológica competitiva.

Los diferentes métodos para elaborar mapas de C&T pueden ser agrupados en dos grandes áreas: los enfoques basados en expertos y aquellos basados en recursos computacionales (Kostoff et al., 1998). Existen de la misma manera, dos tipos principales de técnicas basadas en procedimientos de cómputo para desarrollar relaciones estructurales de la ciencia y la tecnología. Por un lado, los de tipo ‘histórico’ que rastrean el linaje de citación de publicaciones y patentes para elaborar mapas de relaciones que ilustran la evolución temporal de un campo, vgr.: Narin (1989). Por otro lado, aquéllos que explotan el uso del fenómeno de co-ocurrencia; estos últimos se basan en la suposición de que la ocurrencia simultánea de una entidad representa algún tipo de relación. De esta manera, la frecuencia de co-ocurrencias es empleada para cuantificar diversos niveles de asociación que permiten la construcción de mapas que muestran la estructura de los campos de investigación. Los métodos más comunes de co-ocurrencia incluyen los análisis de co-citación (Healey et al., 1986), co-ocurrencia de palabras (Callon et al., 1991, Callon et al., 1979, Callon et al., 1986) y co-nominación (Georghiou et al., 1988).

El origen del análisis de co-ocurrencia de palabras puede ser encontrado varias décadas atrás en los trabajos lexicográficos de Hornby (1942) para explicar la co-ocurrencia de conocimiento, y en la investigación lingüística de De Saussure (1949) describiendo la correlación entre la afinidad de unidades lingüísticas y su aparición en el lenguaje.<sup>3</sup> El desarrollo moderno de este método con fines de evaluación de la investigación se originó hacia finales de los setentas con la propuesta de Callon et al (1979); ésta se basaba en la suposición de que las palabras clave asignadas por el autor para describir a un artículo son seleccionadas para guiar al lector hacia una dirección específica. Esto quiere decir que el autor enrolla al lector en un ‘embudo de intereses’ respecto a la transformación del campo científico en el que ambos trabajan. Sin embargo, para lograr esto, tiene que hacer uso de un conjunto de términos clave socialmente aceptados dentro de la comunidad científica involucrada.

A partir de esta base, cada par de palabras  $i$  y  $j$  que aparecen simultáneamente en un artículo es contada como una co-ocurrencia y el número total de éstas en una base de datos particular es empleada para construir una matriz, en la que  $c_{ij}$  denota el número de co-ocurrencias entre  $i$  y  $j$ , y  $c_{ij} = c_i$  corresponde a la frecuencia de ocurrencia de la palabra clave  $i$ . Tras la aplicación de filtros para eliminar términos no relevantes, la matriz de co-ocurrencia es usada como base para construir un mapa —mediante técnicas gráficas derivadas de métodos de análisis estadístico,<sup>4</sup> que ilustra la estructura e intensidad de las relaciones entre palabras y consecuentemente, entre las áreas de investigación dentro de un campo particular.

A partir de la década de los ochentas, el clásico análisis de Callon ha sido mejorado y extendido, tanto desde el punto de vista de las fuentes o unidades de información que constituyen la base para realizar el análisis, así como desde la perspectiva de los métodos estadísticos empleados para procesar los datos. Así, el primer enfoque ha considerado como fuentes desde palabras incluidas en el título de los artículos (Leydesdorff, 1989) hasta ‘frases’ o más bien conjuntos (de uno, dos y

<sup>3</sup> Hornby, AS, Gatenby, EV & Wakefield, H (1942) *Idiomatic and syntactic English dictionary*, Kaitakusha, Tokyo; De Saussure, F (1949) *Cours de Linguistique Generale*, 4eme Edition, Librairie Payot, Paris; según cita (Kostoff et al., 1998).

<sup>4</sup> Por ejemplo, escalamiento multidimensional.

tres términos) identificados en el texto completo de los artículos —denominada en ocasiones ‘tomografía de bases de datos’ (Kostoff et al., 1998). Por otra parte, la segunda aproximación ha incluido desde la visión pionera concerniente a diagramas de ‘densidad v.s. centralidad’ (Callon et al., 1991), hasta el enfoque de conglomerados que utiliza técnicas de escalamiento multidimensional (Peters and Van Raan, 1993).

Más allá de la identificación de una cartografía de la investigación, el valor de la métrica científica se vuelve relevante cuando contribuye a la toma de decisiones durante la planeación estratégica o durante el diseño de políticas. Desde una perspectiva microeconómica, ésta debe ser vista como una pero no la única, entre las varias etapas requeridas para construir un sistema de información estratégica sobre eventos externos de C&T.

### **Minería de datos y el análisis semántico latente (ASL)**

Ya hemos mencionado que dentro de los estudios lingüísticos cuantitativos —cuyo origen puede rastrearse hasta el siglo XIX (Magerman et al., 2010), el análisis de co-ocurrencia de palabras tuvo un lugar predominante en la evaluación de la investigación científica. La extensión de éste hacia el análisis de texto completo de grandes conjuntos de documentos, fue posible con la disponibilidad de grandes bases de datos textuales en forma electrónica y naturalmente con el creciente poder de los recursos de cómputo. Leopold et al (2004) han hecho una amplia revisión sobre las bases de la minería de textos en el campo de la ciencia y la tecnología mientras que Porter y Newman (2004) acuñaron el término de ‘minería tecnológica’ para referirse a la minería de textos de colecciones de documentos de patentes referidas a un tema específico.

Los métodos de minería de datos se dividen de manera general en supervisados y no supervisados, mismos que son aplicados a documentos de texto que son representados como vectores numéricos que pueden ser procesados con procedimientos de estimación estadística. Los primeros parten de un ‘conjunto de datos de entrenamiento’ —conjunto de instancias o documentos de texto que han sido previamente etiquetados de acuerdo a la clase a la que corresponden, de esta manera, el sistema aprende un modelo que le permite clasificar documentos de acuerdo a las categorías provistas por los datos de entrenamiento. Por otra parte, los métodos no supervisados extraen los patrones de interés directamente de los datos sin la necesidad de hacer uso de datos de entrenamiento. Algunas de las técnicas características de estimación empleadas en este segundo grupo son los de análisis de conglomerados —en los que los ítems son agrupados en función de su similitud o distancia,<sup>5</sup> y la reducción dimensional —basada en la disminución de un largo número de variables en un número menor de constructos que capturan las propiedades principales de los datos.<sup>6</sup>

Dentro de los métodos no supervisados uno de los procedimientos comunes en la recuperación de información es el del modelo del espacio vectorial (MEV), que constituye una representación algebraica de los documentos de texto. El espacio vectorial de una colección de textos se construye mediante la representación de cada documento como un vector que contiene las frecuencias de ocurrencia de las palabras encontradas en el documento. El conjunto de los vectores de todos los documentos integra una matriz términos–documentos que representa a la colección completa. Las relaciones entre los documentos se derivan como ya hemos mencionado,

<http://congreso.investigacion.unam.mx>

<sup>5</sup> Empleando por ejemplo, distancia Euclideana o el coseno de los ángulos entre vectores.

<sup>6</sup> Con técnicas estadísticas tales como el análisis de componentes principales, el análisis factorial y la descomposición en valores singulares.

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax 52 (55) 5616.03.08

División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

mediante el cálculo de similitudes o distancias entre los vectores. La codificación de documentos en vectores es lo que se llama indexación, es decir la constitución de un vocabulario global con identificadores únicos para cada palabra de la colección de documentos. Para mejorar el rendimiento de este proceso, se emplean diversos tratamientos previos que incluyen entre otros: la eliminación de términos que no contribuyen al significado distintivo y contenido de los documentos (artículos, preposiciones, etc.); la reducción de los términos a su 'rama' (común denominador de palabras que comparten el mismo significado, vgr.: 'produc' para producto, producción, producir, etc.);<sup>7</sup> y diversas técnicas de ponderación para identificar de manera más adecuada la naturaleza distintiva de los términos en un contexto dado.

El ASL fue sugerido hacia fines de la década de los ochentas, como una teoría y método para mejorar la recuperación de información<sup>8</sup>, en un esfuerzo por resolver dificultades existentes en el modelo del espacio vectorial. Los principales problemas que se enfrentaban eran el de la sinonimia y el de la polisemia. En el primer caso, en el lenguaje existen diversos términos para referirse al mismo objeto o ente, por lo que en el MEV aparecerán como distantes aunque exista relación entre ellos; y en el segundo, diversas palabras tienen más de un solo significado, conduciendo a que términos no relacionados aparezcan como cercanos en el MEV. Los métodos de recuperación empleados sufrían por tanto de problemas de precisión —porción de ítems seleccionados en los que el sistema acierta (debido a la polisemia), y de llamada —porción de los ítems objetivo que el sistema selecciona (debido a la sinonimia).

Frente a esta problemática se propuso un método cuya idea fundamental consistía en reemplazar la representación en la que un conjunto de términos —que en sí mismo constituye una evidencia incompleta y poco confiable de la relevancia de un documento dado, por otro conjunto de entidades que establezcan indicadores más confiables. Para revelar tales relaciones se emplea la estructura de orden superior o latente en la asociación de términos y documentos (Deerwester et al., 1990). El enfoque modela simultáneamente la relación entre documentos con base en sus términos constitutivos y la relación entre palabras con base en su ocurrencia en los documentos. Al emplear un número menor de dimensiones para la representación con respecto al número de palabras únicas, el ASL induce similitudes entre términos que son útiles para resolver los problemas de recuperación de información mencionados anteriormente. De manera sintética, el ASL es una técnica de aprendizaje no supervisado, basado en procesos estadísticos automatizados. Inicia con una amplia colección de documentos o corpus, construye una matriz de términos–documentos e intenta revelar estructuras de similitud que sean útiles para problemas de análisis de texto y recuperación de información (Dumais, 2004).

El análisis semántico latente se lleva a cabo en cuatro etapas principales, dos de las cuales son también empleadas en el método de modelos de espacios vectoriales; el tercer paso es el que constituye la diferencia principal del ASL:

1. Construcción de una matriz de términos–documentos a partir de un extenso corpus, en la que las líneas corresponden a las palabras y las columnas a los documentos (o incluso a unidades menores como pasajes o frases). Las celdas de la matriz contienen la frecuencia con que una determinada palabra ocurre en cierto documento.

2. Transformación de la matriz de términos–documentos para convertir las frecuencias de ocurrencia en ponderaciones de las mismas; existen diversos algoritmos para llevar a cabo esta ponderación, cada uno de ellos con diversos resultados de desempeño (Dumais, 1992).
3. Reducción dimensional mediante el empleo de la descomposición en valores singulares (DVS),<sup>9</sup> en la que los valores singulares  $k$  más grandes son retenidos. El resultado consiste en que cada documento y término es representado ahora por un vector de dimensiones  $k$  en el espacio derivado por la DVS.
4. La recuperación se lleva a cabo en el espacio reducido. Las similitudes y disimilitudes o distancias son computadas<sup>10</sup> en el espacio reducido y no en la matriz original, porque ahora tanto los términos como los documentos son representados como vectores en el mismo espacio. Por tanto los cálculos relativos a similitudes término–término, documento–documento y término–documento son mucho más fáciles de realizar.

## DISEÑO DE LA INVESTIGACIÓN

Con base en el problema expuesto en la introducción, relativo al alineamiento de las actividades de investigación científica y de desarrollo tecnológico dentro de un entorno particular (una nación), reflejado en sus publicaciones científicas y sus patentes, formulamos la siguiente hipótesis:

*H<sub>1</sub>: En el país bajo estudio (en este caso, México) existe concordancia entre las categorías temáticas de los documentos de investigación científica publicados y las clases de patentes que son solicitadas.*

*H<sub>0</sub>: En el país bajo estudio (en este caso, México) no existe concordancia entre las categorías temáticas de los documentos de investigación científica publicados y las clases de patentes que son solicitadas.*

El objetivo fundamental del trabajo de investigación consistirá entonces en:

- Examinar y evaluar la conveniencia de emplear métodos derivados de la minería de datos —particularmente la búsqueda y recuperación de información a través de la creación de espacios semánticos, así como de métodos de ITC —específicamente los mapas cognitivos de la investigación científica, para identificar la concordancia entre un corpus de publicaciones y un corpus de patentes.

Para poder alcanzar este objetivo es necesario: (i) recuperar un conjunto representativo de publicaciones científicas y de patentes de un país (en este caso México), durante un periodo determinado; (ii) aplicar un criterio de homologación entre las categorías temáticas de las publicaciones y la clasificación de patentes; (iii) generar y procesar los *corpora* de categorías temáticas homologadas; (iv) elaborar los mapas cognitivos pertinentes, generar el espacio semántico de los *corpora* y llevar a cabo pruebas estadísticas para verificar si existen o no, diferencias significativas entre los términos del espacio semántico que permitan asociarlos con

alguno de los dos corpus de documentos. A continuación se describe brevemente el procedimiento específico seguido.

### Recuperación y homologación de documentos<sup>11</sup>

Para las publicaciones científicas se empleó la base *ISI Web of Knowledge* y se colectaron todas las publicaciones generadas durante el año 2010, en las cuales hubiera participado al menos un autor con dirección registrada en México; se obtuvo un total de 7,526 publicaciones. En el caso de las patentes, se utilizó el motor de búsqueda *Espacenet* de la Oficina Europea de Patentes, y se recuperaron las patentes solicitadas durante el año de 2010, que tuvieran dentro de su número de prioridad la clave 'MX' (correspondiente a México); se obtuvo un total de 878 patentes. Con ellos se construyeron las bases de datos de trabajo para el estudio.

Los datos recuperados tienen estructuras específicas, que permiten su exportación a campos equivalentes dentro de bases de datos; dichos campos se refieren a ítems de interés relativos a los documentos. Por ejemplo, para este estudio exploratorio en el cual deseamos trabajar con un número reducido de términos, nos interesan los campos de *categorías temáticas* de las publicaciones y el de la *clasificación internacional de las patentes* (IPC, por sus siglas en inglés). Éstos fueron seleccionados ya que reflejan con bastante precisión las áreas de conocimiento a las que corresponden los documentos; no obstante, ya que la clasificación internacional de patentes responde a criterios y necesidades diferentes en comparación a los empleados para clasificar documentos de investigación científica, fue necesario buscar un criterio de homologación. Para ello se empleó la tabla de concordancia IPC–tecnología,<sup>12</sup> que fue elaborada con la finalidad de equiparar la información de sectores productivos y tecnología que implica la clasificación de patentes, con un cuerpo homogéneo y actualizado de áreas de tecnología que permitiera realizar comparaciones internacionales en diversos tipos de estudios (ver Tabla 1).

Tabla 1. Ejemplo de concordancias IPC–tecnología

CLASE DE PATENTE (IPC)	SUBCLASE DE PATENTE (IPC)	CONCORDANCIA CON SECTOR TECNOLÓGICO	CONCORDANCIA CON CAMPO TECNOLÓGICO
Medical or veterinary science; hygiene	Preparations for medical, dental, or toilet purposes	Chemistry	Pharmaceuticals
Computing; calculating; counting	Data processing systems or methods, specially adapted for administrative, commercial, financial, managerial, supervisory or forecasting purposes.	Electrical engineering	IT methods for management
Heating; ranges; ventilating	Production or use of heat not otherwise provided for	Mechanical engineering	Thermal processes and apparatus

De esta manera, se extrajeron finalmente los campos de categorías temáticas de la base de datos de publicaciones y de concordancias IPC–tecnología de la base de datos de patentes.

<sup>11</sup> Los bancos de información empleados fueron seleccionados por la calidad de su contenido, así como por comodidad de acceso y características del formato en que los datos pueden ser recuperados.

<sup>12</sup> (WIPO, 2008), ver también (Schmoch, 2008), respecto a los criterios empleados para establecer las concordancias.

Adicionalmente, para reducir las dimensiones de las matrices que se generarían posteriormente así como el tiempo de los procesos de cómputo, se seleccionó una muestra aleatoria de 500 campos de cada una de las bases.<sup>13</sup>

### Procesamiento

Tras la extracción y selección de la muestra se obtuvo una matriz de dos columnas, la primera para la identificación del origen del documento (literatura o patentes) y la segunda con el contenido de los respectivos campos de categorías temáticas y concordancias IPC–tecnología. Subsecuentemente, cada línea de la segunda columna de la matriz fue exportada a un documento de texto individual, para integrar los respectivos corpus de documentos de publicaciones y patentes.

Con estos dos *corpora* se procedió a construir una matriz de términos–documentos<sup>14</sup> (en este caso una matriz de co–ocurrencia de 123 términos en 951 documentos<sup>15</sup>) que fue inicialmente transformada<sup>16</sup> empleando una ponderación local multiplicada por una global ( $[\log(m_{ij} + 1)] * \text{Entropía}^{17}$ ) para proceder posteriormente a realizar el análisis de semántica latente y la generación del espacio semántico.<sup>18</sup> Éste quedó finalmente constituido por una matriz, de 951 documentos y 123 términos en las líneas y 97 dimensiones en las columnas. Las líneas de la matriz correspondientes a los documentos fueron finalmente empleadas como ‘observaciones’ para el análisis de varianza.

Adicionalmente, y con la finalidad de realizar algunas comparaciones entre los métodos de ITC y los resultados del análisis de varianza, se elaboraron mapas de la estructura cognitiva de publicaciones, de patentes y de publicaciones y patentes, esta vez empleando los registros totales extraídos, así como las distancias entre los documentos a partir del espacio semántico.

### RESULTADOS

Tomando en consideración en primer lugar, los resultados del simple análisis de ITC, los Gráficos 1 y 2 muestran los mapas de densidad de co–ocurrencia de términos en los corpus completos e independientes de publicaciones y patentes. Como vimos anteriormente, éstos revelan a grandes rasgos la estructura cognitiva de la investigación científica y de la investigación orientada a las aplicaciones industriales; es decir, nos muestran una visión general de las áreas en las que se está publicando y patentando. Por motivos de la información que estamos buscando, estas representaciones son muy generales y eliminan los términos de menor relevancia (por *frecuencia* y por sus *interacciones*), pero si nos interesara profundizar en algún área particular, los mapas pueden ser amplificados en ciertas áreas de interés (ver por ejemplo el Gráfico 3).

<sup>13</sup> Para publicaciones: nivel de confianza=95% e intervalo de confianza de 4.2%; para patentes: nivel de confianza=95% e intervalo de confianza de 2.9%

<sup>14</sup> Mediante el empleo del ‘paquete lsa’ desarrollado por Fridolin Wild para el lenguaje y entorno ‘R’.

<sup>15</sup> La reducción de 1000 a 951 documentos se debe a que para construir la matriz se empleó un umbral consistente en una frecuencia global mínima de 5 en la ocurrencia de términos; es decir se eliminaron los términos que en el conjunto total de documentos tuvieran una frecuencia menor a 5 ocurrencias.

<sup>16</sup> Respecto a las transformaciones, ver (Dumais, 1992).

<sup>17</sup> De acuerdo a la definición de Shannon (Shannon, 1948).

<sup>18</sup> Empleando el programa Ucinet (Borgatti et al., 2002).



restringidos a la estructura cognitiva, pero no tenemos información relativa a las interacciones entre los documentos que constituyen los *corpora*. Es aquí en donde encontramos la utilidad de trabajar con el análisis semántico latente, ya que nos ofrece un espacio reducido con datos continuos susceptibles de analizarse estadísticamente ya sea respecto a los términos o a los documentos

La Tabla 3 muestra los resultados parciales del espacio semántico resultado de la descomposición en valores singulares, que como ya mencionamos quedó reducido a 97 dimensiones, de las cuales las primeras 23 explican el 60.7 % de la varianza. Como indicamos anteriormente, las líneas de la matriz correspondientes a los 951 documentos con sus respectivas 23 dimensiones fueron sometidas al análisis de la varianza, cuyo resultado se muestra en el Cuadro 1. Como puede observarse en este último, se identifican dos variables categóricas (Lit y Pat) correspondientes a los *corpora* de documentos de publicaciones y patentes, y la prueba 'F' revela resultados altamente significativos en doce dimensiones ( $\alpha = 0.01$ ) y significativos en dos dimensiones ( $\alpha = 0.05$ ); adicionalmente las pruebas de estadística multivariada resultan también altamente significativas. A partir de estos resultados, podemos inferir que existen diferencias significativas entre nuestras dos variables categóricas, es decir, provienen de dos grupos claramente diferenciados.

Por lo anterior, se infiere que tenemos que aceptar la hipótesis nula relativa a que no existe concordancia entre las categorías temáticas de las publicaciones científicas y la concordancia IPC–tecnología de las patentes. Esto puede confirmarse en los mapas elaborados midiendo la distancia<sup>19</sup> entre los documentos individuales en la matriz reconstruida tras el análisis semántico latente (Gráficos 4–6). Como puede observarse, existen conglomerados dispersos de documentos entre los cuales no existe relación evidente; sólo acercándonos en detalle resulta claro que hay conglomerados de documentos de patentes entre los cuales existen relaciones (Gráficos 5 y 6), cosa que no ocurre salvo raras excepciones entre los documentos de publicaciones (dentro de los

umbrales manejados en los mapas).

Octubre  
Ciudad Universitaria  
México, D.F.

<http://congreso.investiga.fca.unam.mx>

[informacongreso@fca.unam.mx](mailto:informacongreso@fca.unam.mx)

Teléfonos<sup>19</sup> Empleado el coseno del ángulo entre los vectores.

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax 52 (55) 5616.03.08



División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

Gráfico 4. Mapa de distancia entre los documentos en la matriz reconstruida tras el análisis semántico latente

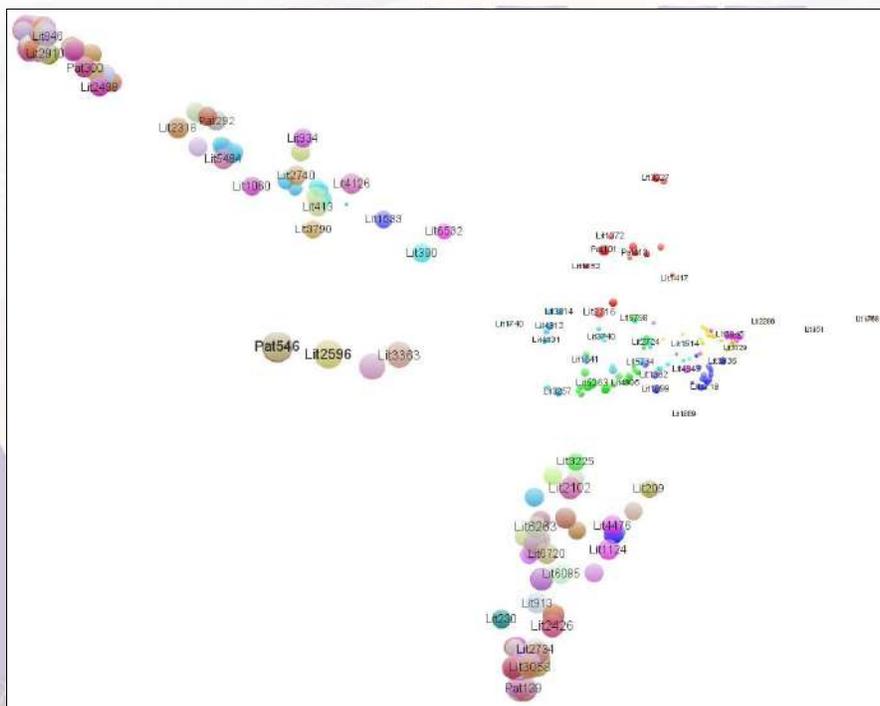
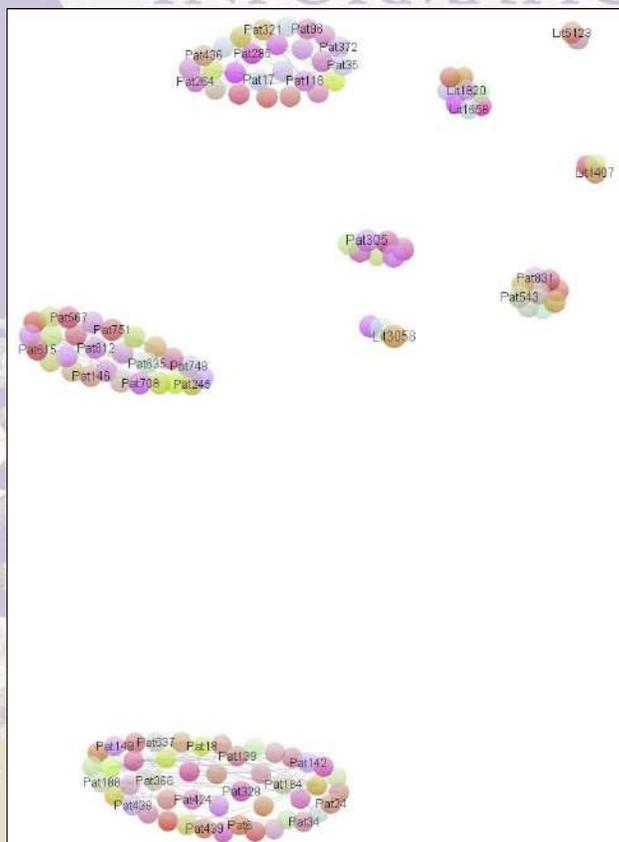


Gráfico 5. Sección inferior central del Gráfico 4



Octubre 5, 6 y 7 de 2011  
Ciudad Universitaria  
México, D.F.

<http://congreso.investiga.fca.unam.mx>  
[informacongreso@fca.unam.mx](mailto:informacongreso@fca.unam.mx)

Teléfonos  
52 (55) 5622.84.90  
52 (55) 5622.84.80  
Fax 52 (55) 5616.03.08

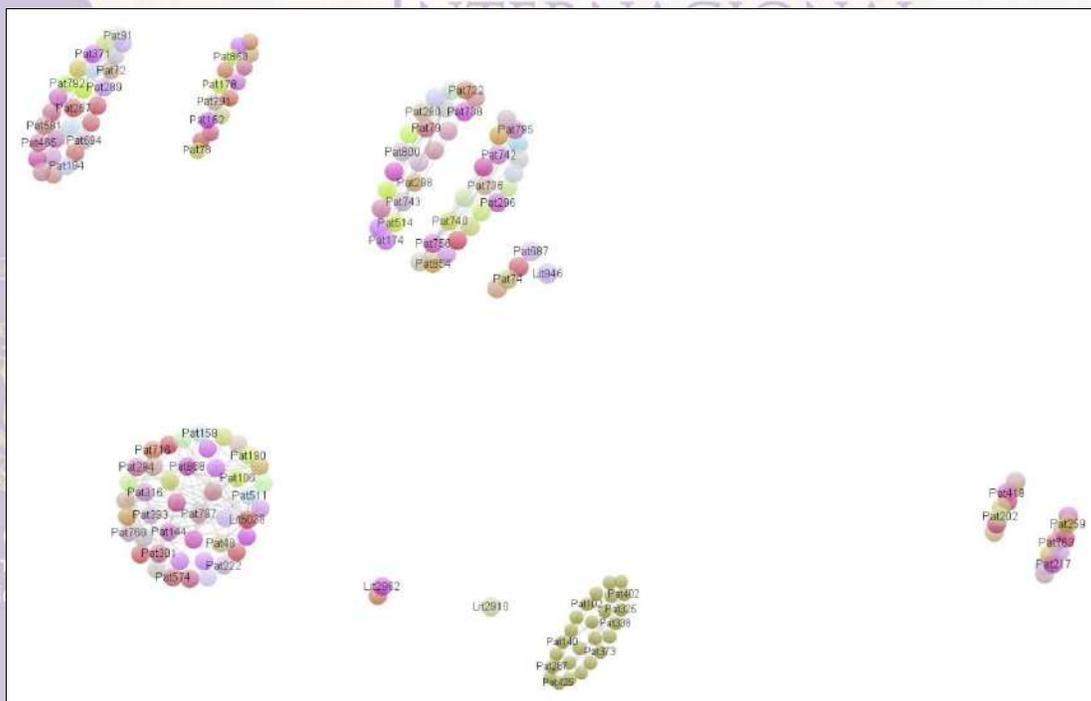


División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

## DISCUSIÓN Y CONCLUSIONES

Hemos visto como el empleo de herramientas de ITC y de minería de datos resultan útiles para analizar grandes colecciones de documentos, en el caso de estudio, publicaciones científicas y patentes. En última instancia, nuestro interés estriba en obtener información útil para la toma de decisión relativa al diseño de políticas de ciencia, tecnología e innovación. Desde esta perspectiva, la falta de alineamiento entre los temas de investigación y los sectores y campos en los cuales se patenta en México, no es del todo sorprendente. Existe una vasta literatura sobre la colaboración academia-industria, que confirmaría el escaso acercamiento entre estas dos esferas. Adicionalmente, habría que mencionar la poca cultura relativa a propiedad intelectual e industrial que existe en el país, que también es evidenciada por el bajo número de patentes identificadas.

Gráfico 6. Sección superior izquierda del Gráfico 4



Tal vez llama un poco la atención, que existan pocas relaciones entre los conglomerados de los propios documentos de publicaciones; la explicación que podemos dar a este fenómeno es que siendo tan vasta la producción de literatura científica, ésta se encuentra mucho más dispersa y por lo mismo es mucho más difícil encontrar conglomerados de publicaciones que muestren interacciones bajo los mismos umbrales que los documentos de patentes. Otro factor que influye en la mayor conectividad entre los documentos de patentes es la homologación realizada mediante la utilización de las concordancias IPC-tecnología, lo que permitió una reducción muy importante del número de términos que normalmente se encuentran en las clases y subclases de patentes.

Desde el punto de vista del análisis de sistemas de innovación, los resultados de estudios de la naturaleza propuesta resultan útiles entonces para ofrecer a los tomadores de decisión datos duros y confiables relativos a algunos de los componentes que constituyen dichos sistemas —el sector de investigación y el sector productivo. De ellos sería posible derivar políticas tendientes a la

reorientación de la investigación, así como otras dirigidas hacia promover el acercamiento del sector productivo con la academia.

Respecto a la perspectiva técnica propuesta, ésta permite hacer más manejable el análisis estadístico requerido para evaluar el alineamiento de los sectores en cuestión. Simplemente en el trabajo experimental realizado y en el cual se seleccionó una muestra relativamente pequeña de documentos, trabajar sobre el espacio vectorial hubiera representado manejar 123 variables en el análisis de varianza. Mientras que el análisis a partir del espacio reducido por el ASL permitió emplear únicamente 23 variables. En síntesis, el método desarrollado es fácilmente sistematizable y sus primeros resultados experimentales permiten confiar en la conveniencia del uso del tipo de herramientas propuestas para análisis de esta naturaleza.

No obstante, aún es necesario extender el alcance de los experimentos con esta técnica, para obtener una confirmación sobre su uso y generalización. Posibles experimentos adicionales incluirían trabajar con otros campos de datos de los documentos tales como los títulos o incluso los resúmenes de las publicaciones y patentes, para contrastar diversos resultados experimentales. Una vez definidos los campos de datos más convenientes para realizar el análisis, podría entonces procederse a realizar estudios comparativos entre países, sobre los cuales existiera información previa sobre sus patrones de publicación y patentamiento, de manera que existiera cierta información de 'control' para la verificación de los resultados.



Octubre 5, 6 y 7 de 2011  
Ciudad Universitaria  
México, D.F.

# XVII CONGRESO INTERNACIONAL DE CONTADURÍA Y ADMINISTRACIÓN E INFORMÁTICA

Comité Organizador: Facultad de Contaduría y Administración, UNAM

<http://congreso.investiga.fca.unam.mx>

[informacongreso@fca.unam.mx](mailto:informacongreso@fca.unam.mx)

Teléfonos

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax

52 (55) 5616.03.08



División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

## REFERENCIAS

- Ashton, W. B. & Klavans, R. A. (1997), "An introduction to technical intelligence in business", en Ashton, W. B. & A., K. R. (Eds.) *Keeping abreast of science and technology: technical intelligence for business*, Columbus, Ohio: Batelle Press.
- Borgatti, S. P., Everett, M. G. & Freeman, L. C. (2002) *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA, Analytic Technologies.
- Callon, M., Courtial, J. P. & Laville, F. (1991), "Co-word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemistry", *Scientometrics*, 22 (1), pp. 155-205.
- Callon, M., Courtial, J. P. & Turner, W. A. (1979) PROXAN: A visual display technique for scientific and technical problem networks. *Second Workshop on the Measurement of R&D Output*. Paris.
- Callon, M., Law, J. & Rip, A. (1986), *Mapping the dynamics of science and technology*, London: MacMillan.
- Deerwester, S., Dumais, S. T., Landauer, T. K., Furnas, G. W. & Harshman, R. A. (1990), "Indexing by latent semantic analysis", *Journal of the Society for Information Science*, 41 (6), pp. 391-407.
- Dumais, S. (1992) Enhancing Performance in Latent Semantic Indexing (LSI) Retrieval. Technical Report. Bell Communications Research.
- (2004), "Latent semantic analysis", *Annual Review of Information Science and Technology*, 38 (1), pp. 188-230.
- Dumais, S. T., Furnas, G. W., Landauer, T. K. & Deerwester, S. (1988) Using latent semantic analysis to improve information retrieval. *Proceedings of CHI'88: Conference on Human Factors in Computing*. New York, ACM.
- Garfield, E., Malin, M. V. & Small, H. (1978), "Citation data as science indicators", en Elkana, Y., Lederberg, J., Merton, R. K., Thackray, A. & Zuckerman, H. (Eds.) *Toward a metric of science: The advent of science indicators*, New York: John Wiley & Sons.
- Georghiou, L., Giusti, W. L., Cameron, H. M. & Gibbons, M. (1988), "The use of co-nomination analysis in the evaluation of collaborative research", en van Raan, A. F. J. (Ed.) *Handbook of quantitative studies of science and technology*, Amsterdam: North Holland.
- Healey, P., Rothman, H. & Hoch, P. (1986), "An experiment in science mapping for research planning", *Research Policy*, 15, pp. 233-251.
- Kostoff, R. N., Eberhart, H. J. & Toothman, D. R. (1998), "Database tomography for technical intelligence: A roadmap of the near-earth space science and technology literature", *Information Processing & Management*, 34 (1), pp. 69-85.
- Leopold, E., May, M. & Paaß, G. (2004), "Data mining and text mining for science & technology research", en Moed, H. F., Glänzel, W. & Schmoch, U. (Eds.) *Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems*, Dordrecht: Kluwer Academic Publishers.
- Leydesdorff, L. (1989), "The Relations Between Qualitative Theory and Scientometric Methods in Science and Technology Studies", *Scientometrics*, 15 (5-6), pp. 333-347.

<http://coninformacongresos.com> Magerman, T., Van Looy, B. & Song, X. Y. (2010), "Exploring the feasibility and accuracy of Latent Semantic Analysis based text mining techniques to detect similarity between patent documents and scientific publications", *Scientometrics*, 82 (2), pp. 289-306.

Teléfonos

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax

52 (55) 5616.03.08

División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

- Molina, A. (1990), "Transputers and transputer-based parallel computers: Sociotechnical constituencies and the build-up of British-European capabilities in information technologies", *Research Policy*, 19, pp. 309-333.
- Narin, F. (1989), "The impact of different modes of research funding", en Evered, D. & Harnett, S. (Eds.) *The evaluation of scientific research*, Chichester: John Wiley.
- Peters, H. P. F. & Van Raan, A. F. J. (1993), "Co-word based science maps of chemical engineering. Part 1: Representations by direct multidimensional scaling", *Research Policy*, 22 (1), pp. 23-45.
- Porter, A. L. & Newman, N. C. (2004), "Patent profiling for competitive advantage", en Moed, H. F., Glänzel, W. & Schmoch, U. (Eds.) *Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems*, Dordrecht: Kluwer Academic Publishers.
- Schmoch, U. (2008) Concept of a Technology Classification for Country Comparisons. Final Report to the World Intellectual Property Organisation (WIPO). Karlsruhe, Germany, Fraunhofer Institute for Systems and Innovation Research.
- Shannon, C. E. (1948), "A Mathematical theory of Communication", *The Bell System Technical Journal*, 27 (July, October), pp. 379-423, 623-656.
- UNCTAD (2011), *Science, Technology and Innovation Policy Review. Perú*, Geneva: United Nations.
- WIPO (2008) IPC-Technology Concordance Table. Geneva, World Intellectual Property Organization, United Nations.

## APÉNDICES

Tabla 2. Palabras con mayor frecuencia en la matriz de términos-documentos

TÉRMINO	FRECUENCIA
chemistry	324
engineering	266
science	137
mechanical	132
physics	79
electrical	79
technology	78
multidisciplinary	77
instruments	76
materials	69
biology	52
civil	51
pharmaceuticals	47
environmental	42
apparatus	40
applied	39
medical	36
food	36

<http://congreso.investiga.fca.unam.mx>

[informacion.congreso@fca.unam.mx](mailto:informacion.congreso@fca.unam.mx)

Teléfonos

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax 52 (55) 5616.03.08

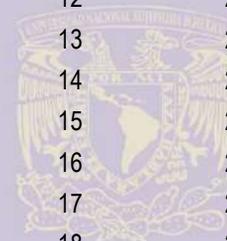


**ANFECA**  
Asociación Nacional de Facultades de Contaduría y Administración

División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

Tabla 3. Resultado de la DVS, primeras 30 dimensiones de una matriz de rango 97

FACTOR	VALUE	PERCENT	CUM%	RATIO	PRE	CUM PRE
1	10.856	7.9	7.9	1.061	0.233	0.233
2	10.235	7.4	15.3	1.987	0.221	0.454
3	5.152	3.7	19.1	1.061	0.055	0.509
4	4.854	3.5	22.6	1.015	0.05	0.559
5	4.78	3.5	26.1	1.268	0.048	0.607
6	3.769	2.7	28.8	1.022	0.03	0.637
7	3.689	2.7	31.5	1.011	0.029	0.666
8	3.649	2.7	34.1	1.079	0.028	0.694
9	3.381	2.5	36.6	1.067	0.024	0.718
10	3.168	2.3	38.9	1.05	0.021	0.74
11	3.018	2.2	41.1	1.078	0.019	0.759
12	2.799	2	43.1	1.052	0.017	0.776
13	2.662	1.9	45.1	1.064	0.015	0.791
14	2.5	1.8	46.9	1.029	0.013	0.804
15	2.43	1.8	48.6	1.063	0.013	0.816
16	2.287	1.7	50.3	1.001	0.011	0.827
17	2.284	1.7	52	1.034	0.011	0.838
18	2.209	1.6	53.6	1.009	0.01	0.849
19	2.189	1.6	55.2	1.028	0.01	0.859
20	2.129	1.5	56.7	1.086	0.01	0.869
21	1.96	1.4	58.1	1.092	0.008	0.877
22	1.795	1.3	59.4	1.019	0.007	0.884
23	1.761	1.3	60.7	1.029	0.007	0.89
24	1.712	1.2	62	1.039	0.006	0.896
25	1.648	1.2	63.2	1.041	0.006	0.902
26	1.583	1.2	64.3	1.037	0.005	0.907
27	1.526	1.1	65.4	1.015	0.005	0.912
28	1.504	1.1	66.5	1.04	0.005	0.917
29	1.447	1.1	67.6	1.053	0.004	0.922
30	1.374	1	68.6	1.012	0.004	0.926



Octubre 5, 6 y 7 de 2010  
Ciudad Universitaria  
México, D.F.

<http://congreso.investiga.fca.unam.mx>

[informacongreso@fca.unam.mx](mailto:informacongreso@fca.unam.mx)

Teléfonos

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax 52 (55) 5616.03.08



**ANFECA**  
Asociación Nacional de Facultades y  
Escuelas de Contaduría y Administración

División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

Cuadro 1. Resultados del análisis de varianza

Variables	Levels					
GROUP\$ (2 levels)	Lit Pat					
N of Cases Processed : 951						
<b>Univariate F Tests</b>						
Source	Type III SS	df	Mean Squares	F-ratio	p-value	
DIM1	50.631	1	50.631	903.638	0.000***	
Error	53.172	949	0.056			
DIM2	6.376	1	6.376	65.111	0.000***	
Error	92.936	949	0.098			
DIM3	1.666	1	1.666	65.508	0.000***	
Error	24.142	949	0.025			
DIM4	1.205	1	1.205	58.896	0.000***	
Error	19.411	949	0.020			
DIM5	0.350	1	0.350	16.399	0.000***	
Error	20.231	949	0.021			
DIM6	0.300	1	0.300	21.993	0.000***	
Error	12.955	949	0.014			
DIM7	0.444	1	0.444	32.241	0.000***	
Error	13.083	949	0.014			
DIM8	0.061	1	0.061	4.463	0.035**	
Error	13.007	949	0.014			
DIM9	0.002	1	0.002	0.152	0.697	
Error	10.979	949	0.012			
DIM10	0.347	1	0.347	35.185	0.000***	
Error	9.372	949	0.010			
DIM11	0.015	1	0.015	1.603	0.206	
Error	9.026	949	0.010			
DIM12	0.106	1	0.106	13.047	0.000***	
Error	7.728	949	0.008			
DIM13	0.005	1	0.005	0.644	0.422	
Error	7.059	949	0.007			
DIM14	0.044	1	0.044	6.936	0.009***	
Error	6.076	949	0.006			
DIM15	0.003	1	0.003	0.415	0.520	
Error	5.903	949	0.006			
DIM16	0.019	1	0.019	3.428	0.064	
Error	5.184	949	0.005			
DIM17	0.124	1	0.124	23.687	0.000***	
Error	4.955	949	0.005			
DIM18	0.111	1	0.111	22.692	0.000***	
Error	4.647	949	0.005			
DIM19	0.004	1	0.004	0.853	0.356	
Error	4.784	949	0.005			
DIM20	0.009	1	0.009	1.818	0.178	
Error	4.522	949	0.005			
DIM21	0.000	1	0.000	0.041	0.840	
Error	3.817	949	0.004			
DIM22	0.000	1	0.000	0.140	0.709	

Oct 6 y 7 de 2011  
 Ciudad Universitaria  
 México, D.F.

http://ci...  
 informac...  
 so@fca.unam.n...

Teléfonos

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax 52 (55) 5616.03.08

División de Investigación, Facultad de Contaduría y Administración, UNAM  
 Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510

**ANFECA**  
 Asociación Nacional de Escuelas de Contaduría y Administración

Error	3.177	949	0.003		
DIM23	0.013	1	0.013	4.137	0.042**
Error	3.085	949	0.003		

**Multivariate Test Statistics**

Statistic	Value	F-ratio	df	p-value
Wilks's Lambda	0.117	304.455	23, 927	0.000***
Pillai Trace	0.883	304.455	23, 927	0.000***
Hotelling-Lawley Trace	7.554	304.455	23, 927	0.000***

\*\*\*  $\alpha = 0.01$

\*\*  $\alpha = 0.05$

# XV CONGRESO INTERNACIONAL DE CONTADURÍA ADMINISTRACIÓN E INFORMÁTICA



Octubre 5, 6 y 7 de 2011  
Ciudad Universitaria  
México, D.F.

Dirección de Comunicación y Relaciones Públicas - Facultad de Contaduría y Administración - UNAM  
 Fotografía: Andrés López Chávez

<http://congreso.investiga.fca.unam.mx>

[informacongreso@fca.unam.mx](mailto:informacongreso@fca.unam.mx)

Teléfonos

52 (55) 5622.84.90

52 (55) 5622.84.80

Fax

52 (55) 5616.03.08



División de Investigación, Facultad de Contaduría y Administración, UNAM  
Circuito Exterior s/n, Ciudad Universitaria, México, D.F., C.P. 04510