

Capítulo 4

Herramientas computacionales para la práctica de la Vigilancia Tecnológica e Inteligencia Competitiva

JENNY MARCELA SÁNCHEZ* – VICTOR A. BUCHELI** – FABIO GONZÁLEZ*

*Universidad Nacional de Colombia, {jmsanchezt, fagonzalez}@unal.edu.co

**Universidad de los Andes, vbucheli@uniandes.edu.co

Resumen

El desarrollo de las organizaciones en entornos altamente cambiantes y globalizados, genera un alto riesgo en los procesos de toma de decisiones. Así, dichos procesos necesitan el apoyo de herramientas que ayuden a orientar la estrategia de la organización y lograr ser competitivos. Una práctica que permite orientar la estrategia de la organización es la Vigilancia Tecnológica e Inteligencia Competitiva (VTIC). Ésta, es un proceso organizado, sistemático y continuado en el tiempo que, apoyado por técnicas y herramientas computacionales, permite el monitoreo del entorno, del cual extrae información de alto valor agregado que apoya los procesos de toma de decisiones. El presente capítulo, pretende dar una visión de la VTIC y los tipos de técnicas y herramientas computacionales que apoyan dicho proceso.

Abstract

Decision making is a highly risky process, and it is accentuated by a dynamic and globalized environment. This process requires tools that support organization strategy guiding, in order to compete in a fastly changing environment. Competitive intelligence (CI) is a discipline that helps to reach this target. CI is an organized, systematic and continuous process that, supported by computational tools, scans the environment looking for information that adds value to the decision making process. This chapter provides an overview of CI and its supporting computational tools.

1: Introducción

Quienes toman decisiones en las organizaciones en entornos altamente cambiantes y globalizados, necesitan herramientas que ayuden a orientar la estrategia de la organización. En tales entornos, es necesario contar con información de alto valor agregado para apoyar dichos procesos de toma de decisiones. Así, es posible valerse de las prácticas de la Vigilancia Tecnológica e Inteligencia Competitiva (VTIC), que es un proceso organizado, sistemático y que, continuado en el tiempo, y asistido por técnicas y herramientas computacionales, puede minimizar el riesgo en dichos procesos de toma de decisiones.

El uso de las técnicas y herramientas computacionales es un elemento clave para llevar a cabo prácticas de VTIC, debido a que permiten apoyar dichas prácticas a través del manejo, recuperación, almacenamiento, administración, procesamiento, representación y visualización de la información. De esta forma, se disminuye el tiempo de búsqueda y se facilita el análisis de la información. Algunos estudios previos [17,18] sobre herramientas computacionales que apoyan la VTIC señalan, primordialmente, que: a) hay una gran cantidad de herramientas que apoyan la búsqueda y recolección de información y b) existe una ventana de oportunidad para el desarrollo de herramientas computacionales que apoyen el análisis, la representación y visualización de la información y de igual forma apoyen el proceso completo de VTIC.

En consecuencia, es pertinente identificar las técnicas computacionales que permitan apoyar todo el proceso de VTIC y trabajar especialmente en técnicas de inteligencia artificial tales como KDT (*Knowledge Discovery Text*), KDD (*Knowledge Discovery Data*), y

TM (*Tech Minig*), entre otras, que hagan posible obtener información de alto valor agregado para el apoyo al proceso de VTIC. Así, en este capítulo se propone una clasificación de técnicas computacionales conforme al proceso de VTIC, y se identifican dichas técnicas en software especializado que apoya las prácticas de VTIC. Por último, se expone la herramienta prototipo (VIGTECH) desarrollada en la Universidad Nacional de Colombia.

La organización de este capítulo, es la siguiente: en la Sección 2, se definen los conceptos asociados a la VTIC; la Sección 3, presenta las técnicas y herramientas computacionales que podrían apoyar la VTIC clasificadas según el ciclo de trabajo; en la Sección 4, se presenta la herramienta informática VIGTECH y, finalmente, en la Sección 5 se presentan las perspectivas de trabajo futuro.

2. Vigilancia Tecnológica e Inteligencia Competitiva (VTIC)

Fenómenos como la velocidad del desarrollo técnico de diferentes áreas, la globalización, la desregulación de diferentes sectores, la informatización, las TIC y el desarrollo de sociedades y economías del conocimiento, amenazan considerablemente muchas organizaciones, e incluso países, que tradicionalmente se aislaron de las posibilidades y niveles de la competencia global pero que, pese a ello, se adaptaron, aprendieron e incluso trataron de modificar su entorno. Es necesario tomar decisiones, anticiparse, estar atento a un sinnúmero de variables y señales del entorno que inciden en el comportamiento y supervivencia de la organización.

En consecuencia, una manera para estar atento a los cambios del entorno de forma sistemática es a través de la práctica de la VTIC¹ que, según la norma AENOR UNE166.006 [1], se define como “una actividad que garantiza la supervivencia de las

¹ En la literatura anglosajona, la inteligencia empresarial se suele denominar *Competitor Intelligence*, *Competitive Intelligence* o *Business Intelligence*. La diferencia entre ellas, radica en que la primera se enfoca únicamente en el estudio de los competidores, la segunda incluye el estudio del mercado, los clientes y los proveedores y la tercera incluye los factores PESTEL (Políticos, Económicos, Sociales, Tecnológicos, Ecológicos y Legales) y cualquier factor externo que pueda afectar los objetivos de la organización. También, se conoce con otras acepciones como *Monitoring*, *environmental scanning* o *competitive early warning*. Por su parte, en la literatura francesa se le denomina *Intelligence Économique*, *veille stratégique*. Cuando esta disciplina se ciñe estrictamente al campo científico tecnológico, se conoce en inglés como *Technological Watch* o *Competitive Technical Intelligence* y en francés como *veille technologique*. Estos sintagmas, incluyen las diversas formas de vigilancia.

organizaciones en un mundo donde se hace necesario estar atentos a todo aquello que se presenta en el entorno, máxime cuando éste es inestable, incierto y complejo”. Así, la VTIC es la implementación, en la organización, de un proceso sistemático en el que se capta, analiza y difunde información de diversa índole—económica, tecnológica, política, social, cultural, legislativa— mediante métodos legales, con el ánimo de identificar y anticipar oportunidades o riesgos, para mejorar la formulación y ejecución de la estrategia de las organizaciones.

2.1. Relación de la VT e IC con otros procesos en la organización

El proceso VTIC, se liga tanto con la gestión de la tecnología, como con la generación de innovación en las organizaciones [2, 3, 4]. En ambos casos, la VTIC es un proceso crítico que permite detectar e interpretar las implicaciones de las amenazas y oportunidades que proceden del entorno. Es decir, la VTIC proporciona información que se utiliza en la producción de nuevos productos o en la venta de los productos actuales.

Igualmente, la VTIC se relaciona, estrechamente, con la gestión del conocimiento, en la medida en que ambas tratan información. Sin embargo, la gestión del conocimiento involucra el tratamiento de la información interna de la organización [5]. Es importante señalar, que una vez la VTIC incorpora la información externa a la organización, se hace necesario que los procesos de Gestión del Conocimiento se definan adecuadamente.

De otra parte, la cantidad de información que se produce en el entorno es abrumadora, de tal forma que quien toma decisiones prefiere apalancarse en las opiniones de los expertos que en la información disponible [6]. En consecuencia, uno de los objetivos de la VTIC es presentar la información del entorno de

El término inteligencia, se adoptó gracias a la supremacía que la literatura en inglés tiene hoy. Adicionalmente, el término inteligencia en el mundo anglosajón significa “información para la acción”. Además, para la cultura hispánica el término inteligencia se define como el “conocimiento o acto de entender y comprender las cosas” y para la lengua francesa se define como “la aptitud para adaptarse a una situación”. Por tanto, la inteligencia abarca no sólo la recolección de información, sino también su comprensión para finalmente actuar, por lo que el término es adecuado para denominar el tema que se está tratando. También, vale la pena aclarar que, para algunos autores como Escorsa y Tena, el sintagma Inteligencia Competitiva es la evolución del sintagma Vigilancia Tecnológica. Así, no tiene sentido hacer vigilancia tecnológica si no se toman decisiones y no tiene sentido generar Inteligencia si previamente no se hace una búsqueda concienciada en el entorno. En suma, en castellano se pueden utilizar los dos sintagmas o el sintagma completo. El punto relevante es que ambos son complementarios y no pueden existir el uno sin el otro.

manera estructurada, de tal forma que el proceso de toma de decisiones alcance el balance necesario.

Finalmente, es importante mencionar que la VTIC se solapa y complementa con otras disciplinas tales como la prospectiva (*foresight*), la predicción tecnológica (*forecasting*), el *roadmapping* y la evaluación tecnológica. Esto sucede, puesto que utilizan métodos similares y tienen como fin último la identificación de oportunidades y riesgos para la organización. El marco de trabajo de integración de estas disciplinas y de sus métodos de trabajo, se conoce como *Technology Future Analysis* (TFA) [7].

2.2. El ciclo de trabajo de la VT e IC

Un proceso de VTIC, realizado de modo sistemático, posibilita el desarrollo de competencias en el ámbito de la innovación, la productividad y la competitividad. Como resultado, las empresas, las universidades e instituciones del gobierno de fomento al desarrollo productivo pueden identificar nuevos servicios y productos y dar una respuesta y una mejor adaptación a diferentes entornos.

Ahora bien, los procesos sistemáticos de VTIC en una organización se articulan a través de la definición de ciclos de trabajo, de roles y procedimientos y de infraestructura al interior de la misma. El ciclo de trabajo se basa en cuatro grandes fases: a) planeación en la que se realiza la definición de la temática en la cual se plantea la especificación del tema principal de vigilancia, los subtemas, los objetivos específicos y los términos y sintagmas para las búsquedas [3, 8, 9, 10, 11]; b) búsqueda y recolección de información sobre la temática pre establecida a partir de información estructurada y no estructurada [12, 13, 14]; c) análisis y generación de inteligencia, en la que se elaboran análisis DOFA, perfil de competidores, análisis semántico, análisis cienciométrico, entre otros, y cuyos resultados tienen un proceso de validación con el equipo de expertos [15, 16]; y, d) difusión, en la que se dan a conocer las implicaciones, recomendaciones y conclusiones con base en los resultados y análisis realizados [17, 18].

Es importante señalar que, de acuerdo con el tipo de VTIC, es decir, del aspecto del entorno en el que se desea profundizar, se tiene un conjunto determinado de preguntas y de fuentes. Así pues, si se desea conocer aspectos de mercado, sería usual preguntarse, por ejemplo, acerca de los segmentos de mercado, de las necesidades futuras del mercado, entre otras, y las fuentes de información que proporcionarían tales respuestas serían los directorios de empresas en Internet, los estudios de mercado, las encuestas de satisfacción de los clientes, etc. Si se está vigilando el

entorno regulatorio, es usual preguntarse acerca de las barreras jurídicas, de la normatividad, de los nuevos proyectos de ley, etc. y su fuente serían los Boletines Oficiales de Estado, el Diario Oficial, los anales del Congreso, etc. Si las preguntas se relacionan con la actualidad tecnológica, las fuentes naturales son las bases de datos de artículos científicos y de patentes.

Paralelamente a la fase de planeación, es necesario definir las herramientas que apoyarán las labores de vigilancia. Dichas herramientas, ya sean automatizadas o no, se encargarán de captar, analizar y evaluar tanto el significado de las tendencias de determinadas tecnologías, como el impacto económico y comercial de los movimientos de los competidores. El éxito de un proceso sistemático de VTIC, se basa en la conjunción de la cultura organizacional con una buena estructura de la organización, adecuados procedimientos, herramientas y tecnologías idóneas. [18]

3. Técnicas y herramientas computacionales que apoyan el proceso VTIC

Son varias las técnicas computacionales y las herramientas de software que pueden apoyar las diferentes fases del ciclo de trabajo de VT e IC. Existen varios estudios que realizan evaluación comparativa de software que apoyen la VT e IC [19, 20, 21, 22, 18, 19, 23]. No obstante, tales estudios no descienden con detalle en las técnicas computacionales que conforman las diferentes aplicaciones. Por ello, aquí se profundizará en las mismas, presentando un estudio comparativo de las técnicas computacionales y cómo cada una de estas puede apoyar las diferentes fases del ciclo de trabajo de la VTIC.

3.1. Técnicas computacionales que apoyan el ciclo de trabajo de la VT e IC

Existen diferentes técnicas computacionales que pueden apoyar las diferentes fases del ciclo de trabajo de VTIC, como se presenta en la tabla 1. Específicamente, las categorías con mayor potencial a desarrollar trabajos de investigación son: a) recuperación de información, b) representación de conocimiento y, c) descubrimiento de conocimiento.

Ahora bien, como ya se mencionó, el proceso de VTIC tiene sentido siempre y cuando se dé valor agregado a los hallazgos. Ello sucede, normalmente, en las fases de análisis y organización e inteligencia. Por esta razón, esta sección se concentra en las técnicas computacionales que apoyan estas fases, trabajando principalmente con información textual.

Técnica		Planeación	Búsqueda y Cartera	Análisis c Inteligencia	Difusión
Almacenamiento	Almacenamiento Semicstructurado	✓	✓		
	Bases de datos estructuradas		✓	✓	
	Gestor de proyectos	✓			✓
Recuperación de información	Búsqueda Booleana		✓		
	Búsqueda en lenguaje natural		✓		
	Crawler – Spider		✓		
	Metabuscadores		✓		
Estadística	Ánalysis de series de tiempo			✓	
	Estadística descriptiva			✓	
	Estadística Multivariada			✓	
Representación de conocimiento	Bases de datos de conocimiento			✓	
	Ontologías			✓	
	XML		✓		
Descubrimiento de conocimiento	Ánalysis de Redes Sociales		✓	✓	
	Data-mining			✓	
	Minería de textos			✓	
	Web-Mining			✓	
	Visualización de datos			✓	✓

Tabla 1. Cuadro de técnicas y fases del ciclo de trabajo de la VTIC

Análisis de series de tiempo: Una serie de tiempo, corresponde a una secuencia de datos ordenados medidos en tiempos sucesivos. El análisis de series de tiempo, se refiere al conjunto de métodos computacionales y estadísticos que permiten entender la dinámica subyacente a la generación de la serie de tiempo y predecir valores futuros de la misma [24].

Estadísticas básicas: se centran en describir un fenómeno a través de la frecuencia con que ocurre un determinado objeto relacionado con el fenómeno. Para el caso específico del ámbito científico tecnológico, existe la cienciometría [25] que se liga totalmente con la VTIC y trabaja bajo un enfoque fundamentalmente estadístico teniendo como unidad la contabilidad de diferentes objetos (por ejemplo, las publicaciones, citaciones, etc.). Sin embargo, el problema de las estadísticas es que generan una representación parcializada y no existe una formalidad al momento de construir diferentes indicadores [26].

Estadística multivariada: para el caso del tipo de datos con los que se trabaja recurrentemente en VTIC, los métodos de análisis multivariados tienen como aplicación el análisis de datos textuales. Por ejemplo, el análisis de correspondencias simples, método de las palabras asociada, etc. son métodos de análisis multidimensional de variables nominales. Estos, consisten en buscar las relaciones y conexiones entre las palabras y clasificaciones documentales descritas en las publicaciones científicas, en los títulos de propiedad industrial, en la web, etc. [27,28]. De esta forma, es posible reconocer las temáticas y

problemáticas que están tratando a través de mapas que representan en alguna medida una categoría documental. En tales categorías, un documento particular tiene una posición relativa en tanto existe una relación entre categorías en el espacio documental [29, 30]. Las técnicas recurrentemente utilizadas son PCA (*Principal Component Analysis*), PCO (*Principal Coordinate Analysis*) y MDS (*Multidimensional Scalling*), debido a que el análisis de datos textuales es un problema de alta multidimensionalidad. Lo que se busca, es reducir las características documentales para poder visualizarlas y, para el caso de algunos algoritmos de minería de datos o textual, transformar las entradas para reducir su complejidad [27,28]. Estos análisis, se complementan con glosarios de palabras, concordancias y la selección de un vocabulario más específico a través de tesauros, ontologías, entre otros.

Bases de conocimiento: permiten el almacenamiento, organización y recuperación de conocimiento. El conocimiento, se puede representar de diferentes formas, como: las reglas, la lógica de primer orden y las redes conceptuales [31].

Ontologías: se enmarcan dentro de las técnicas más eficaces para el análisis y la recuperación de la documentación especializada. Sus principales ventajas son la especificidad de sus términos y sus posibles combinaciones en la recuperación de información. Esta técnica, organiza jerárquicamente un conjunto de conceptos sobre un conocimiento específico, que sirven de soporte a diversas aplicaciones. Una ontología, es un cuerpo estructurado de conocimiento que se basa en un modelo de grafos (nodos y relaciones) para proveer información a otras aplicaciones que contienen los procedimientos y son útiles como “sistemas de representación de conocimiento” [32].

Análisis de redes sociales ARS: es una técnica cuantitativa y estructuralista que permite reconocer los sistemas de relaciones presentes en una comunidad, identificando, de una parte, características de los actores, tales como actores centrales, periféricos, de paso obligado, etc. y, de otra, comportamientos, patrones de enlaces y uniones estables que dan cuenta de vínculos, que expresan la existencia de uniones internas fuertemente ligadas, en las cuales se presentan normas, valores y orientaciones propias de la comunidad [33, 34]. Así, el ARS utiliza elementos tomados del álgebra lineal al igual que de la teoría de grafos para construir desde un conjunto delimitado de actores vinculados entre sí, una representación de las relaciones existentes, al identificar métricas de las estructuras relationales tales como grado, centralidad, cliques, cohesión y patrones estructurales (por ejemplo diádicas, tríadas, hoyos estructurales, grupos o clústeres,

entre otros). Al representar las relaciones existentes en una comunidad con un grafo, es posible evaluar la posición de los actores en la estructura relacional y evaluar la estructura de relaciones [33, 34, 35].

Data mining: la minería de datos, es una herramienta que permite el estudio de conjuntos de datos de gran tamaño para el análisis y descubrimiento de conocimiento [27]. Para el caso del proceso de VTIC, dicha técnica se utiliza principalmente para apoyar la toma de decisiones. Las técnicas de *KDT* (*Knowledge discovery text*) y *KDD* (*knowledge discovery data*) permiten reconocer patrones a través de un proceso sistemático que es útil para identificar las dinámicas de un fenómeno y su evolución para prever cambios y oportunidades [6]. Las técnicas más utilizadas en el proceso de VTIC son las no supervisadas, que buscan principalmente grupos entre objetos similares, partiendo de la idea de representar un documento como un vector de elementos textuales [27, 36] y aplicando una función de distancia que permite identificar la cercanía de los elementos que describe el vector. Con este fin, se desarrollaron diferentes tipos de algoritmos (aglomerativos, de partición y de grafos de partición) que buscan agrupar a los elementos más similares y separar los grupos obtenidos, con el fin de discriminar el conjunto de datos. De esta forma, las técnicas de agrupamiento (*clustering*) son útiles para la exploración de datos, encontrando patrones y características en el corpus documental que no se conocían previamente. El *KDT* (*knowledge discovery in text*) y, específicamente, las técnicas de agrupamiento de textos (*text clustering*), se aplica al proceso de VTIC en el ámbito científico tecnológico, buscando encontrar temáticas a través de las redes de palabras clave de los documentos científicos o de las categorías de los títulos de propiedad intelectual, permitiendo así identificar cuerpos útiles de conocimiento, clasificaciones de los documentos, palabras centrales, dinámicas en la producción de conocimiento, áreas emergentes, entre otros [6].

Text Mining: El objetivo de la minería de textos es extraer información de calidad a partir de información textual. En este contexto, 'información de calidad' implica una combinación de relevancia, interés y novedad. La minería de textos, se basa en técnicas de procesamiento de texto, análisis de lenguaje natural y minería de datos. La minería de textos se aplica desde problemas de extracción de información específica en textos, hasta problemas que impliquen entender el contenido semántico de los documentos [37, 38]. Un área que desarrollada en esta vía, es el *Tech Minig* (TM) [6], que se preocupa por la exploración y explotación de la información científico tecnológica,

permitiendo la comprensión de los procesos tecnológicos o de la innovación. Las fases del proceso de TM son las siguientes: selección de fuentes de información, recuperación de información altamente relevante, limpieza de datos, análisis básicos, construcción de indicadores, mapas, representación de los resultados, interpretación y aplicación de los resultados obtenidos. Así, el TM permite realizar una gran variedad de análisis, encontrando oportunidades de cambio científico tecnológico, planeación y principalmente dar apoyo a la toma de decisiones. Otras de sus funciones son [6]:

- Pronosticar trayectorias probables del desarrollo para tecnologías que emergen.
- Identificar competidores o colaboradores en el desarrollo de un nuevo producto.
- Descubrir mercados potenciales para los resultados de la ciencia y la tecnología.
- Manejar los riesgos del desarrollo y de la puesta en práctica de las tecnologías desarrolladas.
- Pronosticar la tecnología, anticiparse a las trayectorias futuras del desarrollo de tecnologías.
- Identificar pasos evolutivos que siguen el desarrollo de la tecnología.

Visualización de datos: son técnicas que permiten mostrar grandes volúmenes de datos multivariados, de tal forma que es más fácil entender su estructura y relaciones. Esta técnica, es una herramienta fundamental para el análisis exploratorio de datos y tiene como principal fortaleza el uso de la habilidad humana para reconocer patrones visuales [39].

Web Mining: El término minería web, se refiere al conjunto de técnicas usadas para la extracción de información/conocimiento valioso en la web y se puede dividir en tres grandes grupos: minería de contenido web, minería de la estructura web y minería de uso web. La primera, busca la obtención de información en páginas individuales; en muchos aspectos coincide con la minería de textos, pero también se explotan características propias de las páginas HTML. La segunda, busca patrones interesantes presentes en la compleja red de conexiones de la web. La tercera, encuentra patrones interesantes en la navegación o búsqueda de usuarios individuales o grupos de usuarios [40, 41].

3.2. Aplicaciones de software que apoyan el ciclo de trabajo de la VT e IC

El ejercicio de análisis comparativo sistemático de estas herramientas, iniciado en 2002 por Sánchez y

Palop, indica que, para el apoyo de la VTIC, existen dos categorías de software, que se diferencian en que la primera incluye herramientas que apoyan una o dos fases del ciclo de VTIC, mientras que la segunda incluye herramientas que apoyan todas las fases del ciclo de VTIC. La mayoría de las herramientas de software, pertenece a la primera categoría, pues apoyan con mayor capacidad, eficiencia y eficacia una cualquiera de las fases del ciclo de inteligencia. Por tanto, de lo anterior se deduce que las herramientas de software se especializan en alguna etapa del ciclo de VTIC. De otra parte, aquellas herramientas de software que pretenden cubrir todas las fases, y que conforman la segunda categoría, son pocas, e igualmente se observa que son más fuertes en alguna etapa. Ahora bien, según se mencionó anteriormente, es conveniente descender en el análisis de las herramientas de software consideradas en los estudios mencionados. Así, la técnica computacional más utilizada es el *crawler* apoyando la fase de búsqueda y recolección, seguida de las bases de datos estructuradas (véase la tabla 2). Por otra parte, las técnicas detalladas en la sección 3.1 son menos utilizadas, lo que reitera una oportunidad para las casas de software de suplir esta necesidad del mercado. No obstante lo anterior, es importante señalar que, en las técnicas que se destacan en la categoría de Estadística, están las estadísticas descriptivas, en la categoría de Representación del conocimiento, está el uso de XML, y en la categoría de Descubrimiento de conocimiento, están las técnicas de visualización y las de minería de textos.

Un análisis de correlación de las técnicas computacionales usadas en las herramientas de software presentadas (véase la figura 1) señala dos grandes *clusters* de técnicas: *Cluster A*: Gestor de proyectos, *Crawler*, bases de datos estructuradas, almacenamiento semiestructurado, Reportes y visualización, *Knowledge Base*, y metabuscadores. *Cluster B*: *Data o Text o web Mining*, Estadísticas descriptivas, estadísticas multivariadas, ARS, y XML. Estas últimas se dedican a la fase de análisis e inteligencia.

Por su parte, el análisis de correlación de las herramientas de software (véase la figura 2), presenta tres agrupaciones principales: aquellas herramientas que apoyan la fase de búsqueda y captura de información, aquellas herramientas que apoyan la fase de análisis y organización de los hallazgos a través de estadísticas (análisis cienciométricos) y de análisis de redes sociales, y aquellas herramientas que apoyan todas las fases del ciclo de VTIC. Las herramientas que no se agrupan, apoyan una fase del ciclo. Lo anterior coincide con el análisis funcional de cada herramienta presentado por [17] y [18].

4. Herramienta informática VIGTECH

La herramienta informática VIGTECH [42, 43], que a continuación se describe, es una respuesta a las inquietudes que se expresaron antes, como quiera que puede satisfacer la necesidad tanto del mercado de ofrecer herramientas que apoyen las fases de análisis e inteligencia, a través de técnicas relacionadas con la Inteligencia Artificial, como al entorno colombiano en la medida que se desarrolló bajo un esquema de software libre. Así pues, la herramienta informática VIGTECH es un instrumento para facilitar las prácticas de vigilancia tecnológica en una organización, apoyando las fases de vigilancia tecnológica de captación y búsqueda; análisis y organización e inteligencia. La herramienta de software desarrollada, permite el estudio de comunidades científicas, encontrando las relaciones cognitivas y sociales en un conjunto de documentos extraídos de una base referencial tal como ISI. Específicamente, la herramienta soporta las actividades de obtención de información de documentos científicos, extracción de metadatos, cálculo de estadísticas descriptivas, análisis de redes sociales, análisis de redes de palabras claves y visualización.

La herramienta informática VIGTECH utiliza técnicas de aprendizaje de máquina y de minería de datos apoyando así la fase de inteligencia [17], utilizando algoritmos para análisis de redes sociales [33, 34, 35], reducción de dimensionalidad, escalamiento multidimensional [28, 31], agrupamiento [36, 37, 38], entre otros. Estas técnicas, permiten vincular de una forma inteligible los resultados obtenidos, presentando indicadores, mapas, sociogramas y, en general, representaciones relacionales de un tópico dado. La herramienta, se encuentra disponible en la dirección <http://168.176.36.9/~vigtech/>

5. Perspectivas de trabajo futuro

A partir de lo expuesto en este trabajo, es pertinente expandir el estudio de las técnicas computacionales aquí detalladas a otros ámbitos diferentes al de la Ciencia y la Tecnología.

La herramienta VIGTECH que se presenta, la construcción del prototipo, la validación y evaluación por los expertos, demostraron la viabilidad de la idea de construir un primer elemento para el desarrollo de un sistema para la vigilancia tecnológica. Esto, allana el camino para desarrollar investigaciones en esta vía que permitan estudiar las dinámicas y las formas de

producción de conocimiento, dado que las organizaciones que se desarrollan en la sociedad del conocimiento basan sus procesos de producción en éste y, por lo tanto, el estudio del mismo es de total importancia.

	Funciones del sistema informático																			
	Fuentes de datos secundarios		Fuentes de datos primarios		Búsqueda en bases de datos		Búsqueda en bases de datos naturales		Sistemas Spiders		Bibliotecas		Análisis de series de tiempo		Visualización de resultados		Redes de distribución de conocimiento		Análisis	
Australia																				
Autodesk																				
CIS Spider																				
canavita.com																				
Cognos Inc.																				
Dynacore Solutions																				
Dynacore																				
Educational Webware																				
Fetch Agent																				
GlobeTrotter																				
Hootsuite Online																				
Intelligent Star																				
Knowleges Webia																				
Lithium-Analytic																				
MarketTrek																				
Miner																				
OpenPortail																				
PatentINet																				
Patentes Schmidt																				
Patentes Schmid																				
Sector Profesional																				
Strategic Partner																				
Techneagent																				
TextAnalyst																				
Vantageworks																				
VigilWeb																				
Webbarret																				
Websektor																				
Wimote																				
WiseColl																				
Yasca																				

Tabla 2. Cuadro de técnicas y fases del ciclo de trabajo de la VTIC

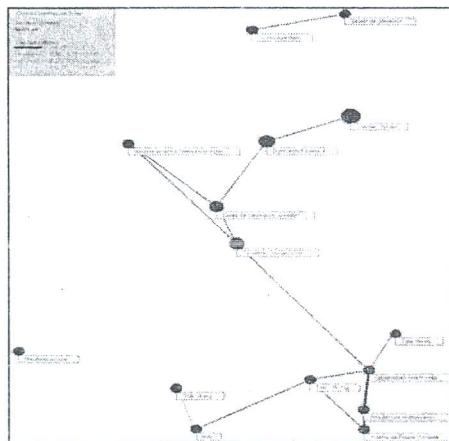


Figura 1. Mapa de correlación de las técnicas computacionales.

De otra parte, el avance en técnicas computacionales de recuperación de información, de minería en texto y de aprendizaje de máquina, permiten encontrar en este primer desarrollo una herramienta que es un paso en la investigación del

estudio a través de otras fuentes de información científica tales como la web, sistemas de registro de patentes, artículos en revistas, reconocimiento de patrones en las dinámicas de producción de conocimiento, evolución de política científica a través de técnicas de aprendizaje de máquina (visibilidad, producción, redes científicas globales, entre otras), desarrollo de sistemas de gestión del conocimiento y uso de tesauros y ontologías para obtener una representación más rigurosa de las representaciones del conocimiento producido por los científicos.

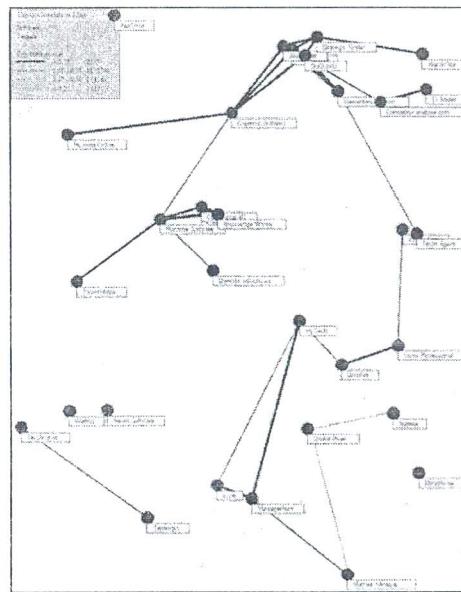


Figura 2. Correlación de las herramientas computacionales.

7. Referencias

- [1] AENOR, "UNE 166.006Ex: Gestión de la I+D+i: Sistema de Vigilancia Tecnológica", Madrid, mayo 2006.
- [2] J. Morin J., R. Seurat, *Gestión de los recursos tecnológicos*, Fundación Cotec. Madrid. 1998.
- [3] P. Escorsa, R. Maspons, *De la vigilancia tecnológica a la inteligencia competitiva*, Ed. Financial Times Prentice Hall, Madrid, 2001.
- [4] P. Escorsa, R. Maspons, I. Ortiz, "La integración entre la gestión del conocimiento y la inteligencia competitiva: la aportación de los mapas tecnológicos", *Revista Espacios* Vol. 21 (2), 2000.
- [5] CETISME, *Inteligencia Económica y Tecnológica. Guía para principiantes y profesionales*, Dirección General de Investigación. Comunidad de Madrid. En: www.madrimasd.org, Madrid, 2003.
- [6] A. Porter, W. Cunningham, *Tech Mining: Exploiting New Technologies for Competitive Advantage*, Wley, New York, 2004

- [7] A. Porter *et al.*, "Technology Futures Analysis Methods Working Group Technology futures analysis: toward integration of the field & new methods", *Technological Forecasting and Social Change*, Vol. 71, No. 3, 2004, pp. 287-303
- [8] L. Fuld, *The New Competitor Intelligence: The Complete Resource for Finding, Analyzing, and using Information about Your Competitors*, John Wiley & Sons, 1995.
- [9] K. Sawka, D. Francis. J. Herring "Evaluating Business Intelligence Systems: How Does your company rate?", *Competitive Intelligence Review*, Vol. 7, Supplement 1, pp. S65-S68, 1996.
- [10] W. Ashton, R. Klavans, *Keeping abreast of Science and Technology. Technical Intelligence for Business*, Batelle Press. Columbus, 1997.
- [11] McDonald, Richardson, *Desingning and implementing Technological Intelligence Systems*, ASHTON y KLAVANS Op Cit. 1997.
- [12] F. Palop, J. Vicente, *Vigilancia Tecnológica e inteligencia competitiva. Su potencial para la empresa española*, Serie Estudios Cotec, No. 15, Fundación COTEC, Madrid, 1999.
- [13] J. Postigo, *Competitive Intelligence in Spain: An investigation into current practices and future possibilities*, Tesina de Master, 2000.
- [14] B. Gilad, *Early Warning: Using Competitive Intelligence to Anticipate Market Shifts*, American Management Association, New York, 2004.
- [15] F. Jacobikak, "Décisions stratégiques et IE: un mariage de raison". *European Symposium 2005 Competia*. Barcelona. Abril. 2005
- [16] A. Porter. "Tech Mining Process: Tecnologías de Análisis de Futuro, Inteligencia Competitiva y Evaluación de Políticas Públicas en Ciencia, Tecnología e Innovación", *Colciencias*, Bogotá, Abril 2006.
- [17] J. Sánchez-Torres, F. Palop, *Herramientas de Software para la práctica de la Inteligencia Competitiva en la empresa*, Triz xxi, Valencia, 2002.
- [18] J. Sánchez-Torres, *Vigilancia Tecnológica e Inteligencia Competitiva*, J. MEDINA y E. ORTEGON Manual de prospectiva y decisión estratégica, CEPAL, 2007, Disponible en www.cepal.org/ilpes
- [19] L. Fuld, *Intelligence Software Report™2000. A Review of Twelve Software offerings in the Competitive Intelligence Arena*, Ed Fuld Company, Londres, 2000/2001.
- [20] L. Fuld, *Intelligence Software Report™2002. A Review of Twelve Software offerings in the Competitive Intelligence Arena*, Ed Fuld Company, Londres, 2002/2003.
- [21] L. Fuld, *Intelligence Software Report™2004. A Review of Twelve Software offerings in the Competitive Intelligence Arena*, Ed Fuld Company, Londres, 2004/2005.
- [22] L. Fuld, *Intelligence Software Report™2004. A Review of Twelve Software offerings in the Competitive Intelligence Arena*, Ed Fuld Company, Londres, 2006/2007.
- [23] A. Comai, J. Tena, J. Vergara, "Software para la vigilancia tecnológica de patentes: evaluación desde la perspectiva de los usuarios", *El profesional de la información*, 2006, Vol. 15, No. 6, pp. 452-458.
- [24] G. Box, G. Jenkins, *Time Series Analysis, Forecasting and Control*, Holden-Day, Incorporated, 1990.
- [25] M. Albornoz, *El Estado de la Ciencia Principales Indicadores de Ciencia y Tecnología*, RICy, 2003
- [26] L. Leydesdorff, "Various methods for the mapping of science", *Scientometrics*, 1987.
- [27] J. Sons, *Data Mining: Concepts, Models, Methods, and Algorithms*, IEEE Press, 2003
- [28] R. Johnson, D. Wichern, *Applied multivariate statistical analysis*, New Jersey, Prentice Hall, 2002.
- [29] F. Moya, C. Rodriguez, "Visualizing the marrow science.", *JASIST*, Vol. 58, pp. 2167-2179, 2007.
- [30] X. Polanco, "Hypergraph modelling and graph clustering process applied to co-word analysis," *Dans Proceedings of ISSI 2007 - Proceedings of the 11th International Conference of the International Society for Scientometrics and Informetrics*, 2007.
- [31] J. Ullman, *Principles of Database and Knowledge-Base Systems: Volume II: The New Technologies*, WH Freeman & Co. New York, 1990.
- [32] S. Chakrabarti, *Mining the Web Discovery Knowledge from Iptext Data*, Morgan Kaufman, 2003.
- [33] D. Wasserman, K. Faust, *Social Network Analysis: Methods and Applications (Structural Analysis in the Social Sciences)* (Paperback), Mark Granovetter SeriesEditor, 2004.
- [34] J. Jariego, L. Teves, "Encuentro de redes en argentina," *Redes – Revista Hispana para el análisis de redes sociales*, Vol. 2, No. 2, pp. 1-10, 2001.
- [35] C. Chen, "The centrality of pivotal points in the evolution of scientific networks," *IUI '05: Proceedings of the 10th international conference on Intelligent user interfaces*, New York, pp. 98-105, ACM Press, 2005.
- [36] Y. Zhao, G. Karypis, "Criterion functions for document clustering: Experiments and análisis", *Machine Learning*, Vol. 55, pp. 311-331, 2004.
- [37] A. Tan. "Text mining: The state of the art and the challenges," *Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases*, 65-70.
- [38] F. Sebastiani, "Machine learning in automated text categorization," *ACM Comput. Surv.*, Vol. 34, No. 1, pp. 1-47, 2002.
- [39] W. Cleveland, *Visualizing Data*, Hobart Press, 1993.
- [40] R. Kosala, H. Blockeel, "Web mining research: a Surrey," *ACM SIGKDD Explorations Newsletter* 2, No. 1, pp. 1-15, 2000.
- [41] P. Turney, *Mining the Web for Lexical Knowledge to Improve Keyphrase Extraction: Learning from Labeled and Unlabeled Data*, National Research Council, Canada, 2002.
- [42] V. Bucheli, F. González, "Herramienta informática para vigilancia tecnológica –VIGTECH", *Avances en Sistemas e Informática*, 2007.
- [43] V. Bucheli, J. Sánchez-Torres, F. González, "Tech-Minig y Vigilancia Tecnológica, El Desarrollo de la Herramienta Informática –Vigtech", *XIV Congreso Latino Ibero Americano de Investigación de Operaciones (CLAIO 2008)*, Cartagena, 2008.