



Universidad Popular Autónoma del Estado de Puebla
Centro Interdisciplinario de Posgrados
Investigación y Consultoría
Departamento de Ingeniería Industrial y Tecnologías
de Información
Doctorado en Tecnologías de Información y Análisis
de decisiones

Tratamiento automático de noticias empresariales en la Web,
mediante minería textual.

Caso de estudio: Universidades privadas de México

Tesis que para obtener el Grado de Doctor en
Tecnologías de Información y Análisis de decisiones

Presenta
Ismael Esquivel Gámez

UNIVERSIDAD POPULAR AUTÓNOMA DEL ESTADO DE PUEBLA



Doctorado en Tecnologías de Información y Análisis de decisiones

TESIS:

Tratamiento automático de noticias empresariales en la Web, mediante
minería textual.

Caso de estudio: Universidades privadas de México

Que presenta:

Ismael Esquivel Gámez

Director de Tesis: Dr. Héctor Simón Vargas

Mayo de 2007.

RESUMEN

El trabajo de investigación, propone y demuestra la vinculación entre las funciones de inteligencia empresarial y las de comunicación social, para que por medio de la minería textual, se pueda develar de manera global, lo que sucede en un segmento de negocios, en cualquier momento.

Para ello, se construyeron tres sistemas computacionales: El primero se encarga de conectarse diariamente a Internet a los sitios Web de las empresas elegidas y copia las páginas Web a un sitio local donde se convierten en archivos texto, se recortan para dejar solamente el título y primer párrafo y se comparan con textos anteriores para evitar su reprocesamiento. El segundo procesa los archivos textos, extrayendo los elementos primordiales de toda noticia y almacenándolos para su posterior consulta. El tercero, se encarga de proveer inteligencia competitiva a los ejecutivos mediante la consulta a las noticias previamente procesadas, a través de una serie de preguntas previamente almacenadas.

Finalmente, y luego de medir el grado de efectividad de la solución y las potenciales aplicaciones a otras áreas del conocimiento, se concluye que el área de minería de textos es prometedora de un caudal de investigaciones, pues la lingüística computacional, disciplina a la que pertenece, ofrece amplias posibilidades de proyectos en nuestro idioma.

ÍNDICE GENERAL

Capítulo I. MARCO METODOLÓGICO	2
INTRODUCCIÓN.....	2
1.1 MOTIVACIÓN.....	3
1.2 DESCRIPCIÓN DEL PROBLEMA.....	3
1.3 OBJETIVOS DE LA TESIS.....	3
1.3.1 Objetivo general.....	3
1.3.2 Objetivos específicos	3
1.4 ALCANCE DE LA TESIS	4
1.5 ORGANIZACIÓN DEL TRABAJO	5
Capítulo II. LA INTELIGENCIA EMPRESARIAL.....	7
2.1 LA TOMA DE DECISIONES EN LOS DIRECTIVOS	8
2.1.1 Proceso de la toma de decisiones	8
2.2 LA INTELIGENCIA EMPRESARIAL	10
2.2.1 Introducción.....	10
2.2.2 Conceptos básicos.....	11
2.2.3 Inteligencia Empresarial	12
2.2.4 Personas y organizaciones	14
2.2.5 Recursos necesarios.	15
2.2.6 Usos de la IE	15
2.2.7 Productos.....	16
2.2.8 Cambio cultural.....	16
2.3 LA INTELIGENCIA COMPETITIVA.....	18
2.4 PROCESO DE INTELIGENCIA	21
2.4.1 Etapas del Proceso de Inteligencia.....	21
2.4.1.1 Diagnóstico	21
2.4.1.2 Planeación.	22
2.4.1.3 Recolección.	22
2.4.1.4 Análisis de la Información.....	23
2.4.1.5 Disseminación	25
2.4.1.6 Actualización	25
2.4.2 Generación de Inteligencia	26
2.4.2.1 Detección de necesidades.....	27
2.4.2.2 Objetivos	28
2.4.2.3 Recolección de datos.....	29
2.4.2.4 Análisis e interpretación.....	31
2.4.2.5 Disseminación	31
2.5 LAS TECNOLOGÍAS DE INFORMACIÓN EN LA INTELIGENCIA EMPRESARIAL.....	33
2.5.1 Minería de datos.....	34
2.5.2 Minería WEB	36
2.5.2.1 Minería de contenidos	38
2.5.2.2 Minería de estructura	38

2.5.2.3 Minería de uso.....	39
2.5.2.4 Aplicaciones.....	40
CAPÍTULO III. ESTADO DEL ARTE DE LA MINERÍA TEXTUAL	41
3.1 LA MINERÍA TEXTUAL	42
3.1.1 Etapas.....	43
3.1.2 Estrategias	43
3.1.3 Aplicaciones.....	44
3.2 ALGUNAS APLICACIONES TÍPICAS DE MT.....	45
3.2.1 Grupo de trabajo 3. Aplicaciones de minería textual y manejo de conocimientos.	46
3.2.2 Minería textual de Ciencia y Tecnología	47
3.2.3 Minería textual en el campo de la mercadotecnia.....	47
3.2.4 Evaluación de aplicaciones de minería textual	48
3.3 PROYECTOS DE MINERÍA TEXTUAL PARA EL TRATAMIENTO DE NOTICIAS.....	50
3.3.1 Clasificación automática de textos de desastres naturales en México	50
3.3.2 Un sistema para la clasificación de notas periodísticas	51
3.3.3 Un sistema de descubrimiento de tendencias para minería de contenido de web dinámico.	52
3.3.3.1 Identificación del recurso.....	53
3.3.3.2 Preprocesamiento.	54
3.3.3.3 Generalización.	54
3.3.3.4 Análisis	54
3.3.4 Encontrando los principales temas en un documento escrito en Español.	55
3.3.5 Nominalización en títulos: una manera de extraer detalles de documentos.	57
3.3.6 Minería de Noticias: Tendencias, Asociaciones y Desviaciones	59
Capítulo IV. LA NATURALEZA DE LAS NOTICIAS	61
4.1 LA COMUNICACIÓN SOCIAL Y SUS FUNCIONES	62
4.1.1 la comunicación social en las universidades privadas	63
4.1.2 la divulgación de eventos.	64
4.2 LA NATURALEZA DE LAS NOTICIAS ESCRITAS	66
4.2.1 el lead	68
4.2.2 la divulgación de la noticia.	69
4.3 ANÁLISIS DE HECHOS NOTICIOSOS	71
4.3.1 análisis lógico de noticias.	73

CAPÍTULO V. PROPUESTA DE TRATAMIENTO AUTOMATIZADO DE NOTICIAS EMPRESARIALES	77
5.1 LAS NOTICIAS EMPRESARIALES EN LA WEB	78
5.2 SELECCIÓN DE LAS EMPRESAS A ANALIZAR	83
5.3 SELECCIÓN DEL MÉTODO DE TRANSFORMACIÓN	85
5.4 SELECCIÓN DEL ESQUEMA DE TRATAMIENTO	88
5.5 SELECCIÓN DE LOS ELEMENTOS POR PROCESAR	91
5.6 TRATAMIENTO MANUAL DE UNA MUESTRA DE TITULARES DE NOTICIAS	94
5.7 CORPUS UTILIZADO	98
5.8 PROCEDIMIENTO PARA EL TRATAMIENTO DE LAS NOTICIAS	100
5.9 ALGORITMO PROPUESTO (MÁXIMA ABSTRACCIÓN)	108
5.9.1 Consideraciones previas	108
5.9.2 Estructuras del encabezamiento	111
5.9.3 Descripción del algoritmo	113
5.9.4 Ejemplo 1. Noticia de tipo AVO	121
5.9.5 Ejemplo 2. Noticia de tipo OVA	124
5.9.6 Ejemplo 3. Noticia de tipo VOA	127
5.9.7 Ejemplo 4. Noticia de tipo VAO	131
5.10 MANEJO DE RESULTADOS	134
5.10.1 Normalización de los resultados	134
5.10.2 Registro de las noticias	134
5.11 EXPLOTACIÓN DE LOS DATOS	137
5.11.1 Planteamiento de preguntas genéricas de inteligencia competitiva	137
5.11.2 Criterios y comandos para dar respuesta	139
5.11.3 Construcción dinámica de los comandos	143
5.12. ARQUITECTURA GENERAL DEL SISTEMA	147
5.12.1 Módulo de Configuración	147
5.12.2 Módulo de Procesamiento	148
5.12.3 Módulo de Consultas	150
5.12.4 Plataforma de ejecución	151
5.13. INTERFACES DE USUARIO DEL SISTEMA	153
Capítulo VI. CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO	155
6.1 CONCLUSIONES	156
6.2 RECOMENDACIONES	158
6.3 TRABAJO FUTURO	159
6.4 IMPACTO SOCIAL	160
Bibliografía y Referencias	161
Anexos	166

Índice de Figuras

Figura	Descripción	Página
2.1	Dominio de la inteligencia.	12
2.2	Proceso de generación de inteligencia.	27
3.1	Minería textual. Proceso de descubrimiento de patrones.	43
3.2	Método propuesto por Méndez-Torreblanca.	53
4.1	Mapa mental del procedimiento propuesto por el Dr. Colle	75
4.2	Ejemplo del procedimiento del Dr. Colle	76
5.1	Una lista de titulares de los eventos	79
5.2	Una descripción breve y completa de los hechos	79
5.3	Una lista de títulos con una imagen asociada a un texto completo	80
5.4	Una lista de noticias con encabezado, una descripción parcial y un enlace a otra página complementaria.	81
5.5	Una relación de encabezado y descripción completa del evento	82
5.6	Un listado de títulos de eventos con enlace hacia páginas que informan a detalle.	82
5.7	Diagrama ERM	101
5.8	Pasos del preprocesamiento (Lectura)	103
5.9	Pasos del preprocesamiento (Escritura)	105
5.10	Etapas de Post-procesamiento	106
5.11	Registro de los resultados	135
5.12	Registro de los resultados con sustitución	136
5.13	Ejemplo de resultados de las consultas	142
5.14	Estructura y vinculaciones de la tabla BASADAS_EN	143
5.15	Pasos para la consulta	144
5.16	Interacciones en datos del módulo de Configuración	147
5.17	Interacciones en datos del submódulo Lector	149
5.18	Interacciones en datos del submódulo Escritor	149
5.19	Intercambio de datos del módulo de Consultas	151
5.20	Plataforma para la ejecución del sistema	152
5.21	Menú principal del sistema	153
5.22	Interfaces de usuario de Catálogos	154
5.23	Pantallas para asociar verbos y referentes a preguntas	155
5.24	Interfaz de usuario para Pregunta-Eventos	155

Índice de Tablas

Tabla	Descripción	Página
3.1	Estrategias empleadas en la Minería de Textos.	43
4.1	Situaciones adecuadas para divulgar una noticia.	65
5.1	Las universidades del sector privado.	84
5.2	Vista parcial de los resultados	95
5.3	Verbos más representativos	96
5.4	Objetos especiales	96
5.5	Objetos más representativos	97
5.6	Descripción de los archivos de la base de datos	101
5.7	Combinaciones de los elementos de mayor frecuencia	111
5.8	Elementos básicos del hecho noticioso	120
5.9	Elementos de una noticia de estructura A-V-O	123
5.10	Elementos de una noticia de estructura O-V-A	126
5.11	Elementos de una noticia de estructura V-O-A	130
5.12	Elementos de una noticia de estructura V-A-O	133
5.13	Contenido de BASADAS_EN y BASADAS_EN1	145

Capítulo I.

Marco Metodológico

En este capítulo se presenta la necesidad de crear un mecanismo para analizar y describir el contenido de grupos de textos, aplicado a la inteligencia competitiva de un sector industrial. Con base en esta necesidad se plantean los objetivos de investigación de este trabajo; los cuales se enfocan en el diseño de un método de minería de texto que permita usar una representación semántica del contenido de los textos y que en consecuencia logre descubrir patrones descriptivos en los textos que consideren información sobre actores, acciones, objetos, atributos y sus relaciones. Como parte final del capítulo se describe la organización de la tesis y el contenido de los siguientes capítulos.

1.1 Motivación

En la actualidad la empresa para ser competitiva requiere conocer oportunamente lo que sucede en la industria en la cual se desarrolla. Al mismo tiempo, requiere informar a sus diversos públicos los eventos que realiza para fines publicitarios y de relaciones públicas. Para informar y ser informada, la empresa utiliza con mayor frecuencia la Web, por las ventajas económicas y de alcance que representa.

Continuamente la empresa está publicando noticias sobre las actividades que realiza y busca lo que están haciendo aquellas que forman parte del sector donde se desarrolla, están vinculadas con el comportamiento que desean informar, permanecen publicadas más de un día y generalmente son de tamaño reducido.

La lingüística computacional se enfoca principalmente en el diseño de los mecanismos que permitan a las computadoras entender el lenguaje natural, aunque también considera varias tareas relacionadas con el procesamiento de información textual. Algunos ejemplos de estas tareas son la búsqueda de información, la extracción de información y la minería de texto.

El procesamiento automático de las noticias mediante la minería de textos provee el descubrimiento de conocimientos no evidentes que surgen de relacionar el contenido de varias de ellas (Hearst y Kodratoff, 1999). Las técnicas usadas también se aplican en la minería Web, en su área de minería de contenido, para archivos tipo texto y HTML.

Un sistema computacional de minería textual que permita, a partir de las noticias empresariales publicadas en la Web, el desarrollo de la inteligencia competitiva para una mejor toma de decisiones estratégicas, es el motor principal para el presente trabajo.

1.2 Descripción del problema

La minería de texto es el área de investigación más reciente del procesamiento de textos. Su proceso principal consiste de: preprocesamiento y descubrimiento. En la primera etapa, los textos se transforman a algún tipo de representación estructurada o semiestructurada que facilite su posterior análisis, mientras que en la segunda etapa las representaciones intermedias se analizan con el objetivo de descubrir en ellas algunos patrones interesantes.

En este trabajo se propone un método de representación formal del contenido de las noticias y un mecanismo para, a partir de ella, responder a preguntas que generalmente se hacen los ejecutivos con respecto a las empresas competidoras.

1.3 Objetivos de la tesis

1.3.1 Objetivo general

Diseñar un nuevo método de minería de texto, apto para representar el contenido de noticias empresariales, que permita trasladar los descubrimientos del nivel temático a un nivel de mayor detalle y conduzca a responder preguntas propias de la inteligencia competitiva.

1.3.2 Objetivos específicos

1. Documentar las ideas principales sobre la inteligencia empresarial y competitiva, la comunicación social y la publicación de noticias empresariales y sobre todo, el estado del arte que guarda la tecnología de información que se pretende aplicar: La minería textual
2. Revisar el estado del arte que guardan las distintas propuestas de tratamiento automático de textos en dominios generales y específicos.
3. Analizar la naturaleza de las noticias empresariales publicadas en la Web, sus componentes, sus estructuras usadas, sus elementos básicos; todo ello para fundamentar una mejor propuesta

4. Plantear los criterios a seguir en la definición del tipo de texto a procesar, el sector industrial a analizar, las empresas más representativas del mismo, los elementos a tratar de la noticia y el esquema de tratamiento.
5. Diseñar un método para la extracción de elementos básicos a partir de las noticias, que permita un registro estructurado de los hallazgos para una posterior explotación.
6. Diseñar un mecanismo que permita responder a preguntas genéricas desde los acontecimientos publicados por las diversas empresas del sector elegido.
7. Diseñar el sistema computacional para la administración de la función de inteligencia competitiva y la diseminación de dicha inteligencia, al equipo humano responsable de la toma de decisiones estratégicas.

1.4 Alcance de la tesis

El trabajo a desarrollar en el presente estudio, tiene como meta obtener un proceso automatizado aplicable a cualquier sector o área industrial, que sirva a las necesidades de sistemas de inteligencia empresarial y/o competitiva. Este proceso automatizado queda únicamente para fines demostrativos, para lo cual se ha escogido una muestra de noticias de ciertas empresas representativas de un sector económico, dejando la replicabilidad de la solución a otros sectores como trabajo futuro. Igualmente, la posibilidad de aplicar el procedimiento desarrollado a noticias de carácter público, como las mostradas en los diarios y revistas, queda en manos de investigadores que deseen determinar el desempeño del citado procedimiento.

1.5 Organización del trabajo

El resto del documento se organiza de la siguiente manera.

En el capítulo 2 se introduce a la inteligencia empresarial como la herramienta esencial para obtener conocimiento actualizado sobre lo que está ocurriendo en un sector empresarial y particularmente lo que están haciendo los actores principales del mismo. Particularmente se habla de la inteligencia competitiva, como elemento esencial de la inteligencia de negocios, así como las modalidades en tecnología de información aplicadas a la misma.

En el capítulo 3 se presenta una breve revisión del estado del arte de la minería de texto. En ella se discuten las principales tareas de la minería de texto, se explican sus principales métodos, se mencionan algunas de sus aplicaciones típicas. Finalmente, se documentan los alcances logrados en proyectos similares sobre textos noticiosos.

En el capítulo 4 se revisa la comunicación social en las empresas y la naturaleza de las noticias escritas que publican, se presentan además, propuestas para el análisis de los hechos noticiosos que implican su transformación a expresiones formales, mismas que permitirán aplicar técnicas de explotación.

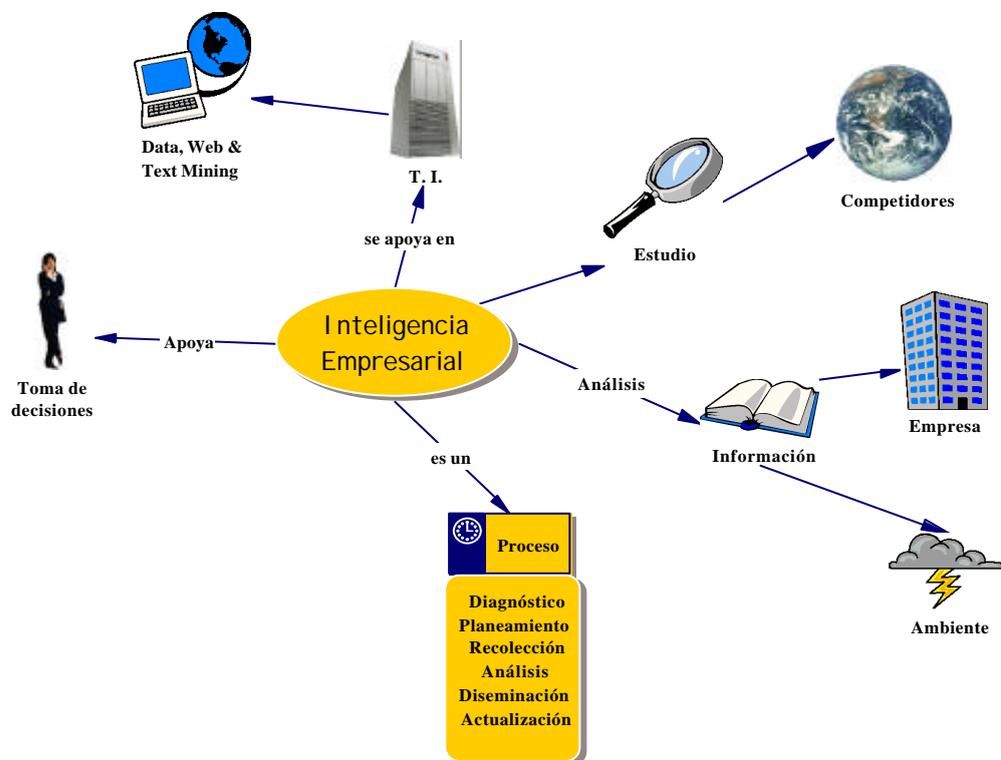
En el capítulo 5 se plantean los criterios usados para la selección de los textos a procesar, el sector industrial por analizar, el método de transformación, el esquema de tratamiento y los elementos de la noticia por procesar. Además, se presenta de manera detallada el procedimiento para el tratamiento y almacenamiento de las noticias y finalmente, se presentan los mecanismos propuestos en la explotación de los elementos encontrados, que permitan responder cuestionamientos desde la perspectiva original planteada: la inteligencia competitiva.

Finalmente, en el capítulo 6 se discuten las principales aportaciones de esta investigación, se mencionan algunos trabajos importantes para el futuro, se realizan las recomendaciones para aplicar el método propuesto en otras instancias y se describe sucintamente el impacto social del trabajo.

Capítulo II.

La inteligencia empresarial

En este capítulo se presentan los conceptos sobre la toma de decisiones, su importancia y el proceso que generalmente se sigue. Se detallan los fundamentos de la inteligencia empresarial, su aplicación en la reducción de incertidumbre, en la reedición de las estrategias y en el monitoreo de los diversos actores del sector industrial en el cual se desarrolla la empresa. Se describen los recursos que se requieren, los usos que se le pueden dar y los productos que pueden obtenerse. Así mismo, se presenta la importancia de un fuerte cambio cultural en la empresa como condición de éxito para un proyecto de inteligencia. Se abordan las características específicas de la inteligencia competitiva, el proceso que se sigue para conseguirla, las técnicas y herramientas más representativas y como prelude del siguiente capítulo, se exponen las tecnologías de información de mayor aplicación.



2.1 LA TOMA DE DECISIONES EN LOS DIRECTIVOS

La toma de decisiones, según Antonio Milano [1], es un proceso metodológico para determinar el curso de acción que mejor satisfaga los objetivos fijado con riesgos aceptables. El proceso es sistemático, racional que ayuda a determinar el curso de acción mejor equilibrado entre la satisfacción de los objetivos perseguidos y los riesgos inherentes. La toma de decisión se realiza posteriormente a un análisis del problema, del cual se conoce la causa para posteriormente mediante este proceso se tome alguna acción.

Una decisión consiste en elegir entre diversas posibilidades, y es el ejecutivo quien hace tal elección. Una decisión puede tomarse de manera inmediata, pero suele suponer un proceso de identificación, análisis, evaluación, elección y planificación. Para llegar a una decisión debe definirse el objetivo, enumerar las opciones disponibles, elegir entre ellas y luego aplicar dicha opción. Las decisiones y el proceso de tomarlas son fundamentales en la gestión. [1]

2.1.1 Proceso de la toma de decisiones

Se formula el propósito básico de la decisión, es decir la meta que se quiere lograr y el resultado final esperado. Esto implica la indicar el objeto y la acción a implantar

OBJETIVOS: Establecer los criterios de selección, los resultados esperados y los recursos con los que se cuenta para lograrlo existen dos tipos de objetivos, los necesarios que son considerados como imprescindibles para el logro del propósito y solo se aceptaran alternativas que cumplan con ellos. Objetivos no imprescindibles pero esperados o también conocidos como deseados y estos son valorados y se les asigna un peso que refleje la importancia que tienen para los que deben tomar la decisión, se le asigna un valor numérico

ALTERNATIVAS: La búsqueda o generación de alternativas será de acuerdo a los objetivos ya planteados, analizando sólo aquellas que cumplan con los objetivos necesarios. Las que sean aceptadas en este proceso se valoran con respecto a los objetivos deseados. Las alternativas que se apeguen a los objetivos deben ser elegidas.

RIESGOS: Se evalúan los riesgos que implica la toma de alguna alternativa, de acuerdo con la información que se tenga disponible, y la experiencia, de acuerdo a la probabilidad de ocurrencia y la gravedad en caso de que llegue a ocurrir

DECISIÓN: De acuerdo al proceso anterior se tiene la información necesaria para analizar, comparar y evaluar las alternativas en forma metódica. Estableciendo cual es la alternativa más equilibrada y que mejor cumpla con los criterios de selección u objetivos y que tenga un grado de riesgo aceptable y que se estén dispuesto a correr.

Ante la necesidad o conveniencia de tomar una decisión, una dirección o un equipo directivo tienen que analizar los diversos elementos o variables más importantes que suelen intervenir:

- a) Responsabilidad de la toma de decisión. ¿Quién decide?
- b) Tipo de decisión. ¿Qué se decide?
- c) Técnica o estrategia para la toma de decisión. ¿Cómo se decide?.
- d) Tiempo. ¿Cuándo se tiene que decidir.
- e) Personas afectadas e implicadas en la toma de decisión.
- f) Consecuencias y repercusiones.

2.2 LA INTELIGENCIA EMPRESARIAL

2.2.1 Introducción

Los profundos cambios acaecidos en la última década sugieren que ya no es posible sobrevivir como empresa, por el mero hecho de trabajar más dentro de los límites internos y menos aún, utilizando prácticas y herramientas tradicionales; la empresa que se limite a trabajar más duro y más rápido, muy probablemente no será competitiva en el mercado global de los años por venir. De acuerdo a Cubillo [2], uno de los elementos esenciales para el éxito empresarial lo constituye el desarrollo de:

Una capacidad de percepción de los factores del ambiente externo, es decir, el desarrollo de mecanismos que permitan detectar y evaluar, con anticipación, oportunidades y amenazas para la empresa; esto incluye por ejemplo, la capacidad para dar respuestas a interrogantes tales: qué están haciendo los competidores y qué son capaces de hacer?; cuáles son las suposiciones o premisas bajo las cuales ellos operan?; probabilidad de nuevos desarrollos tecnológicos y de nuevos productos y su impacto en el sector; nuevos mercados, etc.

Una capacidad de percepción de los factores del ambiente interno tales como: el clima organizacional; la situación financiera y la capacidad de endeudamiento de la empresa; las habilidades y destrezas de los recursos humanos y sus necesidades de entrenamiento, etc., los cuales determinan en gran medida, las debilidades y fortalezas de la empresa.

Una capacidad para anticipar la probabilidad de escenarios, la evolución de problemas y de soluciones asociadas.

En fin, el propósito general es el de reducir la incertidumbre y permitir que la empresa formule un repertorio de maniobras estratégicas ofensivas y defensivas, que le permitan desarrollar y mantener ventajas competitivas en un entorno altamente cambiante.

Ante esta situación se presenta a la inteligencia empresarial (IE) también llamada organizacional o de negocios, como la herramienta fundamental para lograr dicho propósito.

2.2.2 Conceptos básicos

Para Sosa y Cobo [3], previo a la definición de Inteligencia Empresarial, se hace necesario clarificar algunos conceptos que suelen usarse con la libertad que la costumbre de usar sinónimos proporciona. Estos son: dato, información, conocimiento e inteligencia.

DATO: Elementos de conocimiento que carecen de significado por si mismos, cadena de caracteres asociada a un concepto. Su naturaleza es cuantitativa.

INFORMACIÓN: Es un dato o conjunto de datos, elaborado y situado en un contexto, de forma que tiene un significado para alguien en un momento y lugar determinados.

CONOCIMIENTO: Es la información como comprensión. Se halla indisolublemente ligado a la práctica y su difusión e intercambio se realiza a través del lenguaje.

INTELIGENCIA: Es información evaluada y analizada, que se caracteriza por contener elementos de juicio para poder seguir un curso de acción. Posibilita la toma de decisiones porque proporciona un grado de previsión de aquello que puede llegar a causar impacto en la organización.

Las anteriores definiciones permiten encontrar la relación entre información e inteligencia. Dado que la inteligencia se puede definir como el conjunto de habilidades innatas o adquiridas, sumadas a los conocimientos acumulados, que nos permiten interpretar y solucionar los desafíos permanentes de sobrevivencia, se puede encontrar que la relación Información – Inteligencia se define a partir de que la Síntesis y análisis pueden transformar datos en información y que para transformar información en inteligencia es necesario aplicar experiencia y discernimiento.

2.2.3 Inteligencia Empresarial

De acuerdo con Peña [4], es el proceso formalizado por medio del cual un grupo gerencial evalúa la evolución de un sector industrial, las capacidades y el comportamiento de los competidores actuales y potenciales, a fin de mantener o desarrollar una ventaja competitiva. Se encuentra que existe una estrecha vinculación entre un ente y su entorno. Dado a que éste es sumamente dinámico, presenta una gran variedad de oportunidades y amenazas para la sobrevivencia, funcionamiento y desarrollo del ente. El dominio de la inteligencia es, por lo tanto, sumamente amplio y se muestra en la figura 2.1.

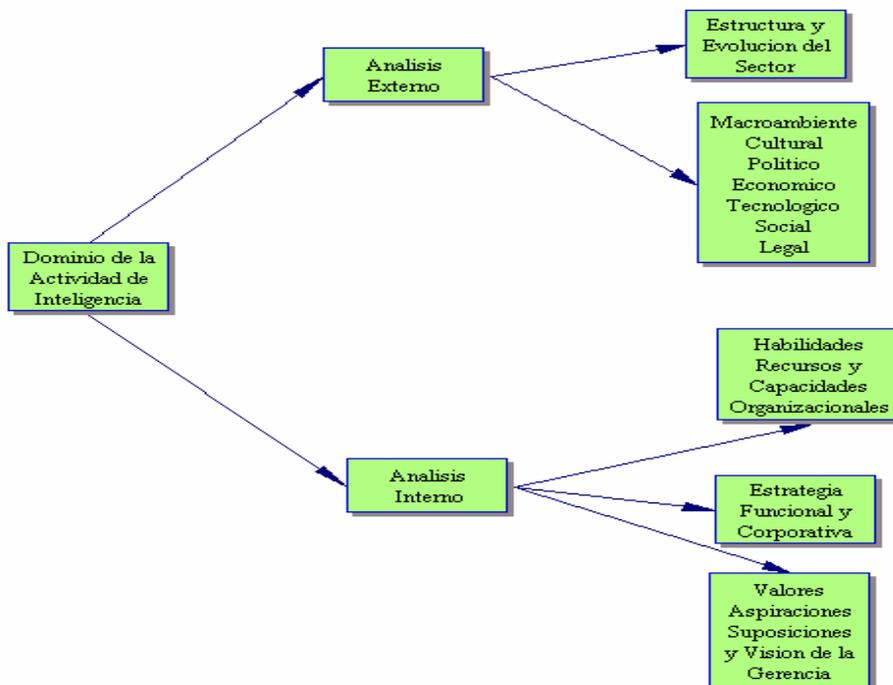


Figura 2.1 Dominio de la inteligencia

Según Sosa y Cobo [3], es la capacidad y función de reunir y analizar datos para obtener y difundir información relevante sobre el ambiente externo y las condiciones internas de la organización de modo sistemático y organizado, que permita crear conocimiento apto para la toma de decisiones y la orientación estratégica.

Arano y Asteggiant [5] refieren que es el proceso a través del cual una sociedad u organización adquiere información en su sentido más amplio, la procesa, la evalúa, la almacena y utiliza para la acción. Involucra un proceso estratégico, y en la actualidad es un recurso de supervivencia ante una sociedad globalizada, donde la inmediatez de las comunicaciones acorta las distancias físicas pero no salva las distancias de desarrollo.

Orozco [6] un destacado autor cubano, estudioso del tema, define con un concepto muy completo a la inteligencia empresarial y precisa que es: una herramienta gerencial cuya función es facilitar a las administraciones el cumplimiento de la misión de sus organizaciones, mediante el análisis de la información relativa a su negocio y su entorno; agrega además que desde el punto de vista del manejo de la información, ella compila, reúne y analiza datos e información, cuyo resultado disemina en la organización, todo lo cual permite obtener de modo sistemático y organizado, información relevante sobre el ambiente externo y las condiciones internas de la organización, para la toma de decisiones y la orientación estratégica. Asimismo describe o prevé hechos y procesos tecnológicos, de mercado, sociales y presenta tendencias. De igual modo usa bases de datos, redes, información de archivos, herramientas informáticas y matemáticas y todo lo necesario para captar, evaluar, validar, analizar información y llegar a conclusiones.

De acuerdo con un trabajo de Susana García [7], define la inteligencia corporativa como la capacidad y la función de reunir, analizar y diseminar datos que permitan obtener, de forma sistemática y organizada, información relevante sobre el ambiente externo y las condiciones internas de la organización para la toma de decisiones y la orientación estratégica.

Por su parte Halal [8], la nombra inteligencia organizacional y se refiere a ella como la capacidad de una organización para crear conocimiento y utilizarlo para adaptarse estratégicamente a su entorno. Otro autor, Tomioka [9], desde su punto de

vista, dice que la inteligencia corporativa es la habilidad de una organización para obtener información del entorno y considera esta última como la fuerza conductora de las organizaciones de autorrenovación.

Torres [10] a partir de las anteriores definiciones establece que: “La inteligencia organizacional es la capacidad de una organización para tomar decisiones efectivas, como resultado del conocimiento adquirido y el conocimiento generado, a partir de la información interna (procedente de los recursos humanos, los procesos, los productos, etc.) e información externa (análisis de tendencias, clientes, competidores). Por tanto, la inteligencia organizacional, es la capacidad intelectual de las organizaciones, que no es precisamente la unión de varias personas inteligentes, soportadas sobre las tecnologías más avanzadas disponibles para realizar sus funciones, sino que en ella, el conocimiento individual se gestiona, comparte y regenera en un nuevo conocimiento de carácter organizacional.”

2.2.4 Personas y organizaciones

Torres [10], llega a establecer que muchas organizaciones se caracterizan por lo que se llama "*crass stupidity*", un término en inglés que caracteriza a organizaciones, que al igual que las personas, se equivocan en la detección de los signos de cambio más obvios, aparecidos en el entorno, en responder apropiadamente a las repetidas demandas de sus clientes, ellas aprenden lentamente y cometen los mismos errores sistemáticamente. Por otra parte, existen organizaciones que muestran las mismas cualidades que se reconocen en las personas inteligentes: curiosidad receptiva y ansiosa, un grupo de respuestas consistentes pero flexibles y la habilidad para aprender rápidamente.

2.2.5 Recursos necesarios.

Para llevarla a cabo la empresa requiere ciertos recursos, entre ellos: Personal preparado en gestión de información y en análisis de información, con conocimiento de los temas de interés a la empresa; acceso a muy variadas fuentes de información, ya sean bases de datos u otras; tecnología de información para el procesamiento más rápido y eficiente; contacto con personas en el entorno informativo de la empresa, ya sea local, nacional o internacional y una clara noción de la gestión de información en función de los intereses de la organización.

2.2.6 Usos de la IE

Planeación Estratégica y de Mercado:

- Perfiles estratégicos (de sector, de compañía, de producto, de personalidad);
- Estudios estratégicos (mercado; introducción de una compañía o producto en una nueva área);
- Estudios de tendencias (técnico, comercial y económico).

Evaluación de Negocios: fusiones, adquisiciones y alianzas;

Investigación y Desarrollo;

Prever Oportunidades y Amenazas;

“Anticipación” de acciones;

Estudios globales de la competencia;

Identificación de potencial oculto en los competidores.

2.2.7 Productos

Los productos típicos de la inteligencia empresarial pueden ser: capacidades de producción de los competidores, planes de desarrollo de nuevos productos, fuentes de nuevas tecnologías, carpeta de patentes, inteligencia para planeación estratégica, identificación de vacíos propios en investigación y desarrollo y tecnología, evaluación de requerimientos tecnológicos para nuevos productos o procesos, identificación de nuevos negocios, identificación de oportunidades comerciales o tecnológicas, perfiles descriptivos de compañía, producto, personalidad. Evaluación de tendencias, Identificación de potencial oculto en los competidores, evaluación de negociaciones y otros, según necesite la empresa y sea capaz de lograrlos mediante el análisis de información obtenida éticamente.

2.2.8 Cambio cultural

De acuerdo con Sosa y Cobo [3], la Inteligencia Empresarial es un modo de:

- Alertar anticipadamente amenazas y oportunidades;
- Monitorear y evaluar movimientos de los competidores;
- Mirar hacia fuera;
- Impulsar el desarrollo de la empresa;
- Estar entre las mejores empresas;
- Apoyar la dirección de la empresa y ser direccionada por ella.

La Inteligencia Empresarial es un nuevo enfoque y una cultura con la que deben contar las organizaciones para el manejo y uso de la información.

Deagostini y Cormenzana [11], refieren que: “Es necesario un cambio cultural que sugiere que la información pertinente puede estar en cualquier parte, y que una de las habilidades a desarrollar consistirá en saber potenciar los procesos comunicativos para poder reconocer la información necesaria en el momento adecuado.” Plantean también que, según Kahaner [12], los japoneses han sabido reconocer a tiempo las virtudes de esta práctica, resumida en los siguientes principios:

- La información tiene siempre un alto valor, ya sea éste intrínseco o práctico.
- La información es ampliamente compartida, inclusive con la competencia.
- La recolección de la información es tarea de todos.

El costo de esta recolección jamás se cuestiona, no se necesita justificar el gasto.

2.3 LA INTELIGENCIA COMPETITIVA

La Inteligencia Competitiva (IC) es una de las disciplinas emergentes que está concitando un interés creciente en el campo de la dirección estratégica. La Sociedad de Profesionales de Inteligencia Competitiva (SCIP) [13] en Estados Unidos la define como un proceso ético y sistemático de recolección de información, análisis y diseminación pertinente, precisa, específica, oportuna, predecible y activa, acerca del ambiente de negocios, de los competidores y de la propia organización. Esta disciplina es el resultado de la integración de algunas áreas del conocimiento. Al ser un campo de investigación reciente son escasos los trabajos que explican sus fundamentos teóricos, a pesar de que las áreas de aplicación actuales han sido múltiples.

Según Walle [14], la inteligencia competitiva (IC) ha emergido como una disciplina independiente, en las ciencias de la administración, pero está muy relacionada con el Marketing. Surgió de la integración de varias perspectivas pues ahora las actividades de inteligencia en las empresas son más integrales, es decir, ahora la orientación no solo es hacia el mercado de clientes, ni competidores, sino hacia atrás en la cadena de valor de las empresas. La IC por tanto, es un proceso. No es espionaje porque no implica prácticas ilegales ni antiéticas de recolección de información y ha sido clasificada como un área dentro de la administración del conocimiento. El competidor inteligente no es IC, pero sí una parte de ella; el valor añadido de la IC está en que toma los conceptos asociados de la Planeación Estratégica y del competidor inteligente.

La IC Se refiere sólo al ambiente de los competidores y sus capacidades, vulnerabilidad e intenciones. Es parte de la inteligencia empresarial, que tiene un enfoque más amplio y abarca todos los aspectos del trabajo de la entidad, incluida la información interna. Es más que estudiar a los competidores, es el proceso de estudiar cualquier cosa que pueda hacer más competitiva a la organización y posicionarla mejor en el mercado. Es parte de la inteligencia empresarial, que tiene un enfoque

más amplio y abarca todos los aspectos del trabajo de la entidad, incluida la información interna.

En términos generales según PUZZLE [15] el ciclo de la inteligencia competitiva se divide en las siguientes etapas:

- **Planificación de las necesidades y definición del contexto de negocio.** Esta tarea se refiere a la definición de las necesidades de los decisores para que puedan alcanzar en mejores condiciones sus objetivos, así como a la definición del proyecto de investigación.
- **Búsqueda y recopilación de la información.** En esta etapa se obtiene la información cuyo análisis y explotación posterior permitirá satisfacer las necesidades de los decisores establecidas en la etapa anterior. Para ello se recurrirá a las fuentes externas e internas disponibles que resulten más adecuadas.
- **Valoración y verificación.** El valor que una información puede tener dependerá no sólo del contenido, sino también de la fiabilidad que ofrezca su fuente y de la posibilidad de contrastarla con fuentes alternativas. En ciertos casos, una información debe ser reservada para uso interno de la empresa, lo que significa que hay que establecer su grado de secreto empresarial.
- **Análisis.** Es el proceso mediante el cual se agrega e interpreta la información y se obtienen las conclusiones y/o recomendaciones que servirán para tomar decisiones.
- **Distribución.** Se refiere a la diseminación de la inteligencia en la organización de acuerdo con los protocolos corporativos o de la empresa.

Estas actividades, se han descrito como una secuencia, en realidad están interconectadas entre sí a través de una retroalimentación continua.

La inteligencia competitiva es un concepto relativamente nuevo, que tiene como objetivos la búsqueda de la “buena” información del entorno externo de la organización, y luego la convierte en un producto inteligente para la toma de decisiones”. Mientras esta actividad, como otras (marketing, I+D) ha adquirido gran importancia durante los últimos años para las empresas, no ha ocurrido lo mismo para las universidades y centros de investigación, en los cuales se halla en etapas incipientes o inexistentes.

2.4 PROCESO DE INTELIGENCIA

En el proceso de la inteligencia empresarial resulta muy importante la veracidad de los datos que se admiten como entrada. Debe destacarse la necesidad de validar y confirmar cualquier información obtenida de fuentes abiertas, inseguras o desconocidas. A continuación se presenta la descripción del proceso de inteligencia de acuerdo a dos estudios:

De acuerdo con Sosa y Cobo [3], se requiere asegurar el proceso de inteligencia empresarial con las capacidades tecnológicas necesarias y los especialistas que tengan las habilidades y el tiempo requeridos para realizar esta función. Para organizar el proceso de obtención y procesamiento de Información utilizamos como herramienta básica destinada a apoyar las decisiones de los niveles estratégicos y tácticos de una organización el Proceso de Inteligencia:

- Diagnóstico = Determinar cuáles son las necesidades de información;
- Planeamiento = Planificación de los procesos a efectuar;
- Recoleccion = Capturar y almacenar la información en bases de datos;
- Análisis = Evaluar y analizar la información;
- Diseminación = Distribución organizada y coherente de resultados;
- Actualización = Evaluación de resultados y readecuación del sistema.

2.4.1 Etapas del Proceso de Inteligencia

Enseguida se presentan cada una de las etapas:

2.4.1.1 Diagnóstico

En esta fase se procede a realizar la Identificación de Necesidades de Información. El desafío consiste en fragmentar el tema en varias cuestiones elementales bastante exactas de forma tal, que la suma de sus respuestas sea la información para el tomador de decisión. Su función es entender cuáles son las necesidades de información del tomador de decisiones y obtener un nivel de

conocimiento tal sobre el tema que permita el comienzo del trabajo. Es una etapa crítica del trabajo. Antes de él, estamos perdidos y sin saber a ciencia cierta qué hacer. Después de él, el trabajo tiene un rumbo a seguir.

2.4.1.2 Planeación.

En esta fase se planea cuidadosamente a partir de las diferentes categorías de información todo el trabajo de monitoreo y recolección de la misma que se va a realizar, teniendo en cuenta que

- Hay varios tipos de información:
 - interna y externa;
 - formal e informal;
 - científica, tecnológica, jurídica, financiera, etc. ;
 - gratis y de pago;

y la pregunta a responder en esta fase es ¿Cuáles fuentes de información se usan y cómo usarlas?

2.4.1.3 Recolección.

Es necesario recolectar toda la información necesaria para llegar a la toma de decisiones. Existen múltiples fuentes de información [Diccionarios, Enciclopedias, Tesoros, Manuales, Indices, Anuarios, Guías, Lista Telefónica, Normas, Patentes, Periódicos, Libros, Bases de Datos, Fabricantes, Fabricantes de equipos, Proveedores, Clientes, Asociaciones, Universidades, Centros de Información, Especialistas, Eventos, Internet (sitios web, buscadores, directorios, metabuscadores, listas de discusión, forums, clippings, bibliotecas virtuales, noticias)]. Una de las fuentes más

modernas para la obtención de información es a través de la INTERNET, en ella se puede localizar:

- Páginas Web: Buscadores, Directorios y Metabuscadore;
- Personas: Catálogos, listas telefónicas y registros
- Archivos: Buscadores de FTP
- Mensajes: Grupos de discusión (Forums) y Listas de discusión
- Noticias: Periódicos, Agencias de Noticias y Clippings
- Informaciones Referenciales: Enciclopedias, Diccionarios, etc.
- Artículos, Patentes e Informes: Bases de datos

Internet tiene fortalezas y debilidades. Por un lado, da acceso directo a gran cantidad de información, sin tardanzas ni costo adicional. Por otra parte, Internet todavía carece de un comprensivo y eficiente sistema de recuperación de información, por lo que los servicios de bases de datos comerciales siguen siendo indispensables para resolver muchos de los problemas asociados al proceso de la inteligencia corporativa. El costo de buscar determinada información en una base de datos comercial puede ser, en algunos casos, menor que el gasto en tiempo dedicado a buscar la misma información en otras fuentes de Internet.

Es necesario establecer un balance entre la localización y selección de las fuentes de información y el empleo de las herramientas existentes para elegir la relevante que permita establecer acciones tendientes a apoyar la toma de decisiones en la organización. La efectividad de las decisiones depende directamente de la disponibilidad y de la calidad de la información al momento de decidir.

2.4.1.4 Análisis de la Información

En la actualidad la cantidad y variedad de información con que se cuenta sobrepasa ampliamente la capacidad humana de asimilarla. Es necesario descartar la información que no aportará elementos al proyecto, y para ello se debe realizar

previamente el tratamiento automatizado de la información para poder efectuar el análisis de la misma. Esta etapa debería ser denominada “generador de inteligencia”. Es en este punto que el analista transforma las informaciones colectadas en una evaluación significativa, completa y confiable. No obstante, debe tenerse presente que nada sustituye la capacidad humana de pensar y reunir todos los ítems en una solución completa para el problema que se está investigando.

Existe gran variedad de herramientas para el tratamiento automatizado de la información y poca integración. Algunas de estas son:

- Inteligencia Competitiva: desarrollo interno, Knowledge.Works (Cipher Systems), STRATEGY! Version 2.5 (Strategy Software, Inc.), Wincite Version 7.0 (Wincite Systems LLC), Strategy! (vide www.fuld.com)
- TextMining: dtsearch, Smart TextMiner, DataSet for MSWord
- DataMining: Excalibur, RetrievalWare, Oracle Darwin, VantagePoint
- DataWarehousing: Oracle Internet BI, Informix I-SPY
- Mapeamiento del Conocimiento: MSEExchange, TheBrain.com
- CRM: Siebel Applications, Intelligent CRM
- Benchmarking: desarrollo interno, benchmarkindex
- GroupWare/Workflow: MSEExchange, Documentun
- Agentes Inteligentes: Copernic, Kengi, TextWise
- Tratamiento y análisis: Dataview, Matrisme, Infotrans, VantagePoint

Métodos de análisis (entre otros):

- DAFO
- Data Mining
- Clasificación automatizada
- Análisis bibliométrico
- Análisis estadístico
- Método Delphi

2.4.1.5 Diseminación

Hacer que la información analizada llegue a las personas que la necesitan de forma útil y agradable. Entre otros ejemplos podemos citar: informes escritos, presentaciones orales, informes personalizados, boletines, etc. Existen dos formas básicas de diseminación:

- Diseminación enfocada o dirigida - Cuando determinado usuario o grupo de usuarios solicita un tipo de información específica;
- Diseminación general - Cuando la inteligencia es diseminada para toda la empresa sin un usuario específico.

Aquí deben considerarse algunos puntos importantes:

- Definición de mecanismos de distribución, responsabilidad de los analistas de información;
- Definición del lenguaje, forma y facilidad de acceso al producto generado por el sistema;
- Definición de la frecuencia de envío de los productos a los clientes;
- Credibilidad del análisis.

2.4.1.6 Actualización

Si el proceso termina en la diseminación, la organización solamente habrá adquirido información, ya que la inteligencia sólo ocurre cuando los resultados del proceso son utilizados en la definición de acciones organizacionales.

Es necesario mantener actualizada toda la información generada.

- Evaluación de los resultados obtenidos y decisión sobre:
 - necesidad de corrección;
 - necesidad de complementación;
 - necesidad de actualización.

- El trabajo realizado puede generar nuevas cuestiones

Según refieren, se hace necesario después de finalizado el proceso de inteligencia que se vuelva a reiniciar el proceso para de esta forma evaluar las nuevas necesidades producto de los nuevos escenarios que han surgido y de esta forma volver a comenzar y reajustar el sistema a la nueva situación existente, esta debe ser nuevamente analizada y diseminada para seguir ejerciendo la inteligencia.

2.4.2 Generación de Inteligencia

Ahora bien, de acuerdo con Peña [4], la generación de inteligencia puede ser vista como un proceso que se inicia con la determinación de necesidades de información de la empresa, el establecimiento del objetivo general del programa, la recolección de información, análisis e interpretación de la misma y la diseminación a los públicos adecuados, tal y como se muestra en la figura 2.2.

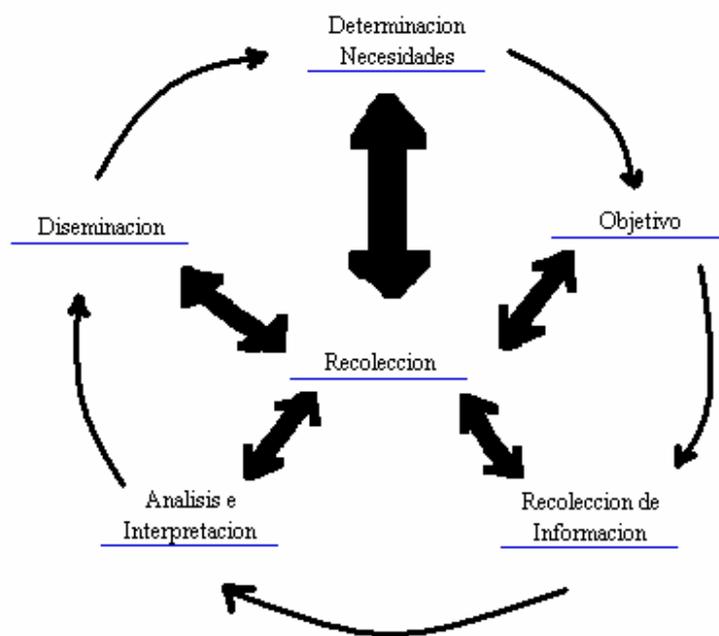


Figura 2.2 Proceso de generación de inteligencia

2.4.2.1 Detección de necesidades.

Para esta se requiere de una auditoría de información. Esta es una etapa crítica pues además de conocer las necesidades de información, se evalúan otros aspectos tales como: hasta qué punto los recursos internos de información satisfacen las necesidades detectadas, cuáles son las prácticas y las actitudes de la gerencia y del personal en relación a las fuentes, el procesamiento y la diseminación de la información en la empresa y cuáles son los canales de comunicación más utilizados.

Según Gilad y Gilad [16], la auditoria de inteligencia debe responder las siguientes preguntas:

- ¿Qué inteligencia es requerida por los usuarios?
- ¿Qué información está disponible?
- ¿Quiénes pueden actuar como recolectores de información?
- ¿Quiénes son los expertos de la empresa en un campo dado?

¿Cuáles son los canales de comunicación en la empresa?

Es importante señalar que el análisis de los resultados de la auditoría de la información constituye el plano detallado según el cual se construirá el programa de inteligencia. Una vez que han sido conocidas las necesidades de inteligencia de la empresa, se hace indispensable establecer cuál o cuáles de ellas deben ser satisfechas primero, es decir se deben establecer prioridades. Según Lagerstam [17], algunos de los criterios utilizados para jerarquizar las necesidades de inteligencia incluyen:

Importancia de tener conocimientos sobre un evento: lo cual es una función de la trascendencia del evento para la empresa, a mayor trascendencia mayor debe ser la prioridad de generar inteligencia; la velocidad con la cual el evento puede impactar la organización, si la velocidad es alta, la empresa tendrá poco tiempo para reaccionar y en consecuencia, la prioridad de generar inteligencia será alta; finalmente también debe considerarse la capacidad de respuesta de la empresa, si esta capacidad es alta entonces la necesidad de inteligencia disminuye.

- Probabilidad de ocurrencia del evento.
- Costo de anticiparse comparado con el costo de reacción en diferentes momentos.

2.4.2.2 Objetivos

El establecimiento de los objetivos es una de las actividades cruciales para el éxito de cualquier tarea de inteligencia. Existen tres tipos de objetivos: ofensivos, defensivos y de reconocimiento y ellos van a caracterizar los proyectos de generación de inteligencia. Los proyectos ofensivos son aquellos que se adelantan cuando se desea evaluar las fortalezas, las debilidades y las posibles respuestas de los competidores que pueden incidir en el éxito o fracaso de un movimiento táctico o estratégico de nuestra empresa. Los proyectos defensivos son aquellos que tienen como propósito anticipar o por

lo menos comprender, los movimientos de los competidores que pueden de una u otra forma amenazar la posición competitiva de la empresa, y a la vez, desarrollar repuestas que neutralicen esas amenazas.

Los proyectos de reconocimiento son aquellos que tienen como propósito conocer mejor el sector o las actividades que desarrollan los competidores. Los resultados de este tipo de proyecto pueden conducir a acciones ofensivas o defensivas.

2.4.2.3 Recolección de datos

La clave de esta etapa está en dirigir todos los esfuerzos de recolección hacia los objetivos de la tarea de inteligencia. Las fuentes de información a ser utilizadas en una tarea de inteligencia, van a depender tanto de la tarea en sí misma como de las limitaciones organizacionales (básicamente tiempo y dinero). Los métodos para la recolección de información pueden ser agrupados según la fuente: bibliotecas y campo. Por biblioteca se refiere a aquellos sitios, públicos o privados, en los cuales material impreso, cintas de video o cintas de audio pueden ser accedidas. Si la información requerida para la tarea de inteligencia no ha sido documentada, el especialista debe entonces recurrir a los métodos de campo que involucran la realización de entrevistas y cuestionarios.

Las fuentes de información también pueden ser clasificadas en fuentes internas y fuentes externas. Las fuentes internas son aquellas que han sido generadas por el personal de la empresa o comisionadas a terceros, entre éstas se pueden señalar:

- Documentos producto de las operaciones internas: Estas fuentes vitales de información están asociadas a las operaciones cotidianas de la empresa. Ellas pueden ser generadas en una base diaria, semanal, mensual, trimestral, etc.; se incluyen aquí informes de producción, informes de ventas, de control de calidad, informes financieros, informes de control y seguimiento de materias primas y suministros, etc.

- Informes analíticos: Estas fuentes de información pueden ser definidas como aquellas que contienen un análisis de las actividades de la empresa o de su entorno. Comparadas con los documentos de operaciones, los informes analíticos son preparados y distribuidos para los ejecutivos. Se incluyen en esta categoría, informes de mercadeo, informes de ventas, estudios de planificación, etc.
- Estudios adquiridos a terceros: Con frecuencia las empresas contratan externamente la elaboración de estudios específicos, por ejemplo estudios de mercado, y aunque son generados externamente son utilizados internamente.
- Manuales: De organización, de procedimientos, de productos, normas técnicas internas, etc.

Las fuentes externas son aquellas que proveen información o inteligencia generada fuera de la empresa. Los gerentes y técnicos en su intento por mantenerse al tanto de las operaciones y del desempeño interno de la empresa, consumen (y generan) una gran cantidad de información; sin embargo, con frecuencia dejan de lado un volumen sustancial de literatura externa que incide directamente sobre el funcionamiento de la organización, a pesar que una porción importante de su toma de decisiones se basa en este tipo de información.

Ante la inmensa cantidad de fuentes, la empresa debe desarrollar un conjunto de criterios que le permita racionalizar la adquisición y el procesamiento de la información generada extramuros; entre estos criterios podemos mencionar:

- Nivel técnico de las publicaciones: divulgativas, informativas o altamente especializadas.
- Cobertura geográfica: si la publicación sólo incluye información sobre mercados y competidores locales o si por el contrario, su cobertura es internacional.

- Cobertura temática: si la publicación suministra información sobre otras industrias o sectores con los cuales está vinculada nuestra empresa.
- Otros criterios a considerar son el idioma en el cual aparece la publicación, la periodicidad (quincenal, mensual, trimestral, etc.) y el costo de suscripción.

2.4.2.4 Análisis e interpretación

Una vez finalizada la recolección de la información, se inicia el proceso de análisis e interpretación de los datos. Esta etapa requiere una gran coordinación pues en ella se integran tanto la información interna como la generada externamente. El primer paso consiste en evaluar la información obtenida, para lo cual deben considerarse aspectos tales como: confiabilidad de la fuente, confiabilidad y relevancia de la información, su actualidad o vigencia y si ésta es o no suficiente. El análisis involucra la prueba de hipótesis, el tratamiento de la información divergente así como el reconocimiento de patrones en la información por medio del uso de métodos estadísticos.

A partir de la información, los especialistas deben realizar sus predicciones, lo cual es una actividad bastante compleja debido a la variedad de factores interrelacionados que influyen en ella. Se entienden como predicciones a las presunciones cultas sobre tendencias y eventos futuros. La mayoría de las empresas tiene cierta experiencia con las técnicas utilizadas, ya que han estimado o realizado proyecciones sobre sus ingresos y utilidades anuales, porcentaje de participación en el mercado, etc.

2.4.2.5 Diseminación

Una vez que la información ha sido digerida y analizada, es indispensable contar con los mecanismos que permitan que la inteligencia generada llegue en forma expedita a todos los usuarios que la requieran. Un sistema de distribución de inteligencia debe considerar el uso de informes (de inteligencia crítica, situacionales y los solicitados por el usuario), de carteleras, salas de demostración, boletines y

reuniones. Si la empresa cuenta con los recursos suficientes, es conveniente contar con una base de datos y una red local de computadoras.

- La frecuencia, el contenido y el formato de los informes variará según la inteligencia a diseminar, así tenemos:
- Los informes de inteligencia crítica son altamente específicos en contenido y descriptivos. Deben ser diseminados rápidamente ya que requieren acción inmediata; su frecuencia es ad hoc. Deben ser comunicados vía teléfono, correo electrónico o por otro medio expedito.
- Los informes de situación son por lo general descriptivos, de cobertura limitada y tienen como propósito informar sobre actividades de los competidores y sobre cambios en el ambiente (interno y externo).
- Los informes solicitados por los usuarios pueden ser amplios o muy precisos, según la solicitud de los mismos; deben ser entregados en la fecha establecida y su frecuencia y formato dependerán de las necesidades del usuario.

2.5 LAS TECNOLOGÍAS DE INFORMACIÓN EN LA INTELIGENCIA EMPRESARIAL.

Desde cualquier punto de vista que se observe, debe entenderse que cuando se define la inteligencia organizacional, como la capacidad para reunir, analizar y diseminar información no solo interna sino también sobre el mercado, los competidores y los proveedores, es necesario procesar grandes volúmenes de información, imposible de realizar sin la creación de una infraestructura tecnológica en la organización que permita procesar, analizar, almacenar y distribuir dicha información. La aplicación de los enfoques de la inteligencia organizacional requiere de redes locales que garanticen el flujo de información en la institución, bases de datos, técnicas y herramientas para el análisis de los datos disponibles, así como de un acceso libre a Internet, como una enorme fuente de información que posibilita la realización de búsquedas a bajo costo y la comunicación interpersonal y grupal, entre otros.

“Uno de los impactos más fuertes de Internet en el interior de las organizaciones es el relacionado con la aplicación de las tecnologías web en los ambientes corporativos. Su migración a las redes locales (intranets) y corporativas (extranets) ha aportado una maravillosa interfaz que permite normalizar la presentación de la información de las organizaciones en una forma gráfica atractiva e independiente de las plataformas de los servidores y de las estaciones de trabajo. Esta posibilidad significó un importante salto en el camino para la creación de una cultura de la información -acceder a la información, compartir la información, gestionar la información y consumir la información requerida- en las empresas.” Vilaragut [18].

Ante esta situación, han existido grandes avances de tres tecnologías de información, las cuales serán descritas sucintamente a continuación:

- Minería de datos (Data Mining)
- Minería Web (Web Mining)
- Minería textual (Text Mining)

2.5.1 Minería de datos

De acuerdo con Estivill-Castro [19], la minería de datos es el análisis de archivos y bitácoras de transacciones con el fin de descubrir patrones, relaciones, reglas, asociaciones o incluso excepciones que sean útiles para la toma de decisiones.

El descubrir patrones o relaciones útiles en una colección de datos ha recibido tradicionalmente muchos nombres. El término *data mining* llegó incluso a ser muy desprestigiado en la estadística, pues representaba masajear suficientemente los datos hasta que los mismos confirmasen lo que uno quería postular. En este sentido, la minería de datos es un proceso que invierte la dinámica del método científico en el siguiente sentido.

En el método científico, primero se formula la hipótesis y luego se diseña el experimento para coleccionar los datos que confirmen o refuten la hipótesis. Si esto se hace con la formalidad adecuada (cuidando cuáles son las variables controladas y cuáles experimentales), se obtiene un nuevo conocimiento.

En la minería de datos, se coleccionan los datos y esperamos que de ellos emerjan hipótesis. Queremos que los datos nos describan o indiquen por qué son como son. La más inocente mirada a los datos por un humano, puede inspirarle una hipótesis. Recuérdese que los humanos tenemos grandes poderes de generalización e identificación de patrones. Luego entonces, validar esa hipótesis inspirada por los datos en los datos mismos, será numéricamente significativa, pero experimentalmente inválida. De ahí que la minería de datos debe presentar un enfoque exploratorio, y no confirmador. Usar la minería de datos para confirmar nuestras hipótesis puede ser peligroso, pues estamos haciendo una inferencia poco válida.

2.5.1.1 Fases de un Proyecto de Minería de Datos

De acuerdo a Infleza [20], los pasos a seguir para la realización de un proyecto de minería de datos son siempre los mismos, independientemente de la técnica específica de extracción de conocimiento usada:

- Filtrado de datos
- Selección de Variables
- Extracción de Conocimiento
- Interpretación y Evaluación

Las prácticas de minería de datos se realizan con base a procedimientos como:

Clasificación. Consiste en examinar las características de una entidad nueva y asignarle una clase predefinida. Por ejemplo: Clasificar a un nuevo cliente según su riesgo de crédito (alto, medio, bajo).

Estimación. Similar a lo anterior, pero aplicado a variables continuas. Por ejemplo: ingresos, balance de tarjetas de crédito, etc..

Predicción. Predicción de fidelidad de clientes "Churn Modeling". Por ejemplo: predecir qué clientes nos abandonarán en los próximos 6 meses

Grupos Afines o Reglas de Asociación. El objetivo de los grupos afines es determinar que cosas van juntas. También puede ser usado para identificar oportunidades de "cross-selling" (venta cruzada de productos complementarios). Por ejemplo: que productos debieran ir juntos en un supermercado, recomendaciones de productos.

Clustering. Clustering tiene como objetivo el segmentar a un grupo diverso en un conjunto de subgrupos o "cluster". A diferencia de clasificación, clustering no depende de clases predefinidas. Y es el primer paso en segmentación de mercado. Por ejemplo: un cluster particular de síntomas puede indicar una enfermedad particular

Descripción y Visualización. Algunas veces el objetivo es simplemente describir qué está ocurriendo en una base de datos compleja, para así aumentar nuestro

entendimiento de las personas, productos o procesos que generaron los datos inicialmente.

2.5.1.2 Aplicaciones

Las prácticas de minería de datos se utilizan en diversas industrias y funciones de negocios.

Empresas de telecomunicaciones, tarjetas de crédito y compañías de seguros utilizan minería de datos para la detección de fraudes, optimización de campañas de marketing, descripción y segmentación de clientes, predicción de fidelidad de clientes (churn modeling).

La industria del comercio utiliza la minería de datos para diseñar y evaluar campañas de marketing, definir ofertas más apropiadas o recomendaciones de productos (cross y up selling) a clientes, y predecir riesgo en asignación de créditos a clientes.

La industria de la medicina utiliza la minería de datos para predecir la efectividad de procedimientos quirúrgicos, exámenes médicos y medicamentos

Bancos e Instituciones Financieras para monitorear las transacciones de sus cuentahabientes y reducir usos indebidos de las tarjetas.

2.5.2 Minería WEB

De acuerdo con Vicente [21], la minería de web se refiere al proceso de descubrir la información o conocimiento potencialmente útil y previamente desconocido a partir de datos web. En la Red, en cuanto a información, no hay estándares, no es estructurada y es heterogénea. Crece y cambia muy rápidamente, y esto hace necesario disponer de un mejor descubrimiento de recursos y de la

extracción de conocimiento. Se pueden descubrir elementos de información interesantes que aportan información y conocimiento valiosos. La minería de web combina comunicación, objetivos y técnicas de diferentes áreas.

Hace notar que se utilizan técnicas de minería de datos para descubrir automáticamente los documentos y servicios del web y extraer información de ellos. La minería de web se descompone en las siguientes tareas:

Descubrimiento de recursos: localización de documentos relevantes o no de la red. Ésta es la función de los índices buscadores (que extraen contenido basándose en palabras, zonas del documento, idioma) y de los índices temáticos (que clasifican los documentos). Extracción de información: extracción de determinada información a partir de un documento, independientemente de su formato (HTML, XML, texto, ps, PDF, LaTeX, FAQ, etc.).

Generalización: descubrimiento de patrones generales a partir de sitios web individuales (agrupamiento de documentos o clustering, asociaciones entre documentos).

Análisis, validación e interpretación de los patrones

Taxonomía de la minería de web.

Minería de contenidos. Consiste en diversos tipos de datos, tales como texto, imágenes, audio, vídeo, metadatos e hiperenlaces. La búsqueda reciente en la minería de múltiples tipos de datos se llama minería de datos multimedia.

Minería de estructura. Intenta descubrir el modelo subyacente de las estructuras de los enlaces del web. El modelo se basa en la topología de los hiperenlaces con o sin la descripción de los enlaces.

Minería de uso. Intenta dar sentido a los datos y comportamientos generados en las sesiones de navegación del web.

Mientras que la minería de contenidos y estructura utiliza los datos reales o primarios del web, la minería de uso mina los datos secundarios derivados de las interacciones del usuario con el web.

Técnicas usadas

2.5.2.1 Minería de contenidos

Text Mining: se trata fundamentalmente de técnicas de recuperación de información (IR), técnicas estadísticas y lingüísticas. Hay dos tipos de aplicaciones: text categorization y text clustering.

Hypertext Mining: no se refiere sólo a los enlaces entre documentos, sino también a enlaces intradocumentos. Se realiza con la ayuda del grafo de referencias.

Markup Mining: las marcas contienen información (las marcas HTML contienen secciones, tablas, negritas, cursivas; las marcas XML contienen aún más información).

Multimedia Mining: campo poco desarrollado, principalmente se realiza sobre bibliotecas de imágenes.

La minería de contenidos puede dividirse en dos grandes apartados:

- Minería de contenidos de páginas web
- Minería de los resultados de búsquedas

2.5.2.2 Minería de estructura

Según Chakrabarti [22], consiste en estudiar la estructura de los enlaces entre documentos e intra documentos. Las técnicas se inspiran en el estudio de las redes sociales y el análisis de citas. Una página (persona o artículo) está reforzada por la

cantidad de referencias (amistades o citas) que tiene. El análisis de la estructura es muy útil a la hora de descubrir:

Autoridades, que proporcionan la mejor fuente sobre un tema determinado.
"Hubs", que proporcionan colecciones de enlaces a autoridades.

La minería de estructura puede dividirse en dos grandes apartados:

Utilizando los enlaces
Utilizando la generalización

2.5.2.3 Minería de uso

Se centra en técnicas que pueden predecir el comportamiento del usuario cuando interacciona con el web.

Esta información puede residir en:

- Clientes web (por ejemplo, cookies)
- Servidores
- Proxies
- Servidores de banners (doubleclick.com...)

Previamente a la minería, hay que procesar la información:

- Eliminar los reintentos
- Separar a los diferentes usuarios
- Unir sesiones diferentes
- Juntar páginas con marcos (frames)
- Filtrar por tiempo
- Extraer las páginas poco relevantes
- Otros

El resultado del preprocesado pueden ser:

- Datos específicos para métodos específicos
- Datos relacionales (una base de datos normal)
- Datos en XML

La minería de uso puede dividirse en dos grupos:

Trazado de patrones de acceso general

Trazado de patrones personalizado

2.5.2.4 Aplicaciones

La nueva generación de buscadores

Si tenemos en cuenta que las búsquedas son la segunda tarea que más se hace en Internet después del correo electrónico, tendremos que ver nuevos productos de búsqueda que intenten reducir la recordación de la información recuperada y poder sugerir al usuario aquello que más le convenga, teniendo en cuenta las búsquedas previas e incluso pensar en nuevas maneras de agrupar la información (los mapeadores).

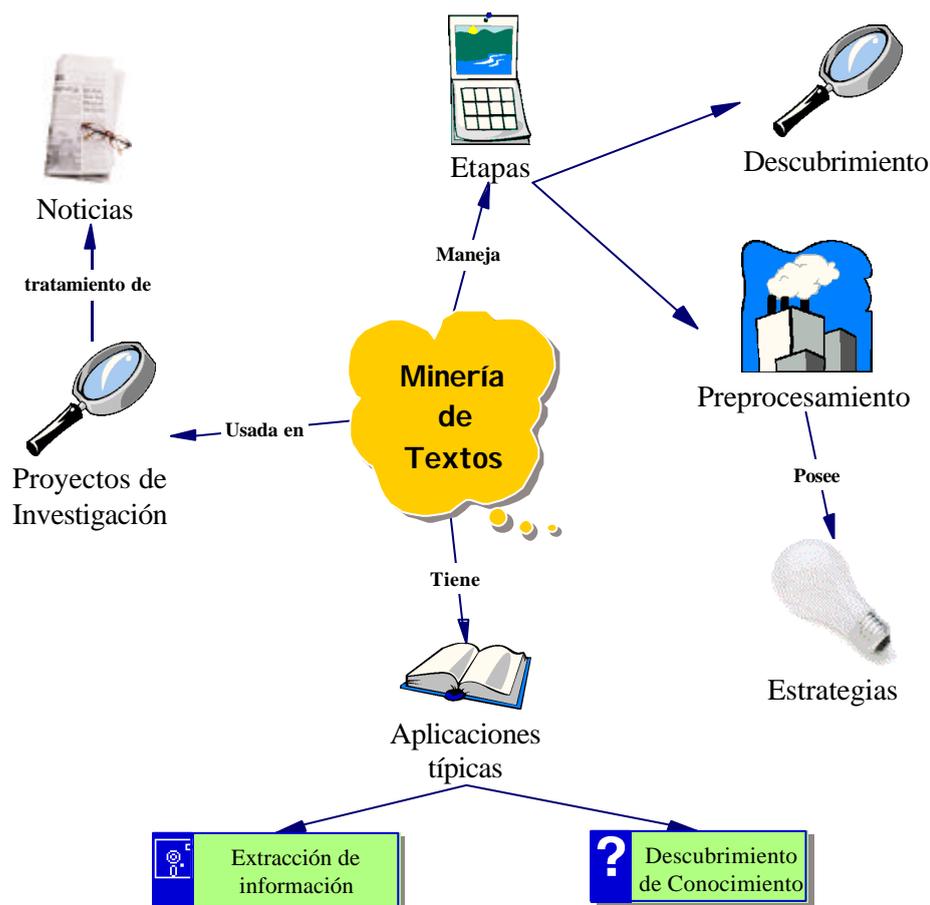
La inteligencia competitiva

Los campos de la documentación donde ello es de fácil e inmediata aplicación es en la inteligencia competitiva, ya que permite realizar el monitoreo y la vigilancia de forma automática.

Capítulo III.

Estado del arte de la Minería Textual

En este capítulo se discuten los fundamentos de la minería de textos, las etapas que se siguen en la transformación, las estrategias por aplicarse dependientes del tipo de representación requerida, los usos más comentados, así como las áreas de conocimiento donde mayormente se aplica. Se presenta además, el estado del arte que guarda la minería textual en cuanto a proyectos de tratamiento de textos sobre hechos noticiosos y se comentan brevemente los resultados que se alcanzaron.



3.1 La Minería Textual

Conforme a Marín et. Al [23], la minería textual es una herramienta para analizar elementos de texto con el fin de identificar, deducir y ampliar conocimiento a partir de cualquier organización de documentos (documentos electrónicos, bases de datos, etc.), puede entender semánticamente el contenido de uno o más documentos y crea una catálogo de los temas tratados. En la práctica, se podría considerar una extensión del tipo de web mining de contenidos. Su objetivo es también la categorización de texto, el procesamiento en lenguaje natural, la extracción y recuperación de información o el aprendizaje automático.

Es una herramienta para analizar elementos de texto con el fin de identificar, deducir y ampliar conocimiento a partir de cualquier organización de documentos (documentos electrónicos, bases de datos, etc.), puede entender semánticamente el contenido de uno o más documentos y crea una catálogo de los temas tratados. En la práctica, se podría considerar una extensión del tipo de web mining de contenidos. Su objetivo es también la categorización de texto, el procesamiento en lenguaje natural, la extracción y recuperación de información o el aprendizaje automático.

Los mismos autores refieren que: “No se debe confundir text mining con la recuperación de información (retrieval information) según la definición “clásica” heredada de los SGD (Sistemas de gestión documental): estos parten de la recuperación de documentos como respuesta a una petición basada en palabras clave y términos de indización dentro de una ecuación de búsqueda, en vocabulario controlado o en texto completo. El término text mining no dispone, en cambio, de ese punto de partida, sino que se refiere a un proceso en que se examina una colección de documentos para descubrir información que no está individualmente en ningún ítem sino que está representada en un determinado segmento de conocimiento producto de una cadena de implicaciones causales”

3.1.1 Etapas

Según Montes [24], la minería textual se define como el proceso de descubrimiento de patrones interesantes y nuevos conocimientos en una colección de textos. Este proceso consiste de dos etapas principales: una etapa de preprocesamiento y una etapa de descubrimiento (Tan [25]). En la primera etapa, los textos se transforman a algún tipo de representación estructurada o semi-estructurada que facilite su posterior análisis, mientras que en la segunda etapa las representaciones intermedias se analizan con el objetivo de descubrir en ellas algunos patrones interesantes o nuevos conocimientos.

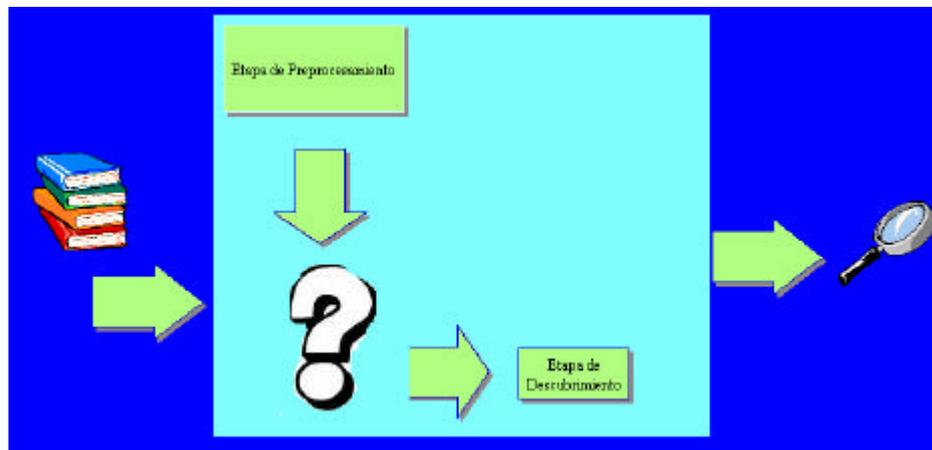


Figura 3.1 Minería textual. Proceso de descubrimiento de patrones.

Dependiendo del tipo de métodos usados en la etapa de pre-procesamiento es el tipo de representación del contenido de los textos construida; y dependiendo de esta representación, es el tipo de patrones descubiertos.

3.1.2 Estrategias

Los tres tipos de estrategias empleadas en los actuales sistemas de minería de texto son:

Etapas de pre-procesamiento	Tipo de representación	Tipo de descubrimientos
Categorización	Vector de temas	Nivel temático
Full-text	Secuencia de palabras	Patrones de lenguaje
Extracción de información	Tabla de datos	Relaciones entre entidades

Tabla 3.1 Estrategias empleadas en la Minería de Textos.

El autor considera que todos estos métodos limitan a un nivel temático o de entidad sus resultados, haciendo imposible descubrir cosas más detalladas como:

- Consensos, que por ejemplo respondan a preguntas como: ¿Cuál es la opinión mayoritaria de los mexicanos sobre el gobierno de Fox?
- Tendencias, que indiquen por ejemplo si han existido variaciones en la postura de Fox con respecto a la educación.
- Desviaciones, que identifiquen por ejemplo opiniones “raras” con respecto al desempeño de la selección mexicana de fútbol.

3.1.3 Aplicaciones

De acuerdo con Bordoni y d'Avanzo [26], las aplicaciones de la minería de textos se usan fundamentalmente para:

- Extraer información relevante de un documento (resumiendo, extrayendo lo más notable, etc.).
- Adquirir perspectivas sobre las tendencias, las relaciones entre gentes/lugares/organizaciones, etc. agregando y comparando automáticamente la información extraída de documentos de un cierto tipo.
- Clasificar y organizar documentos según su contenido; es decir, preseleccionar automáticamente grupos de documentos con un tema específico y asignarlos a la persona adecuada.
- Organizar depósitos de meta-información relacionada con documentos para la búsqueda y recuperación. Recuperar documentos basándose en varios tipos de información sobre el contenido del documento.

La lista de actividades muestra que las principales áreas de aplicación de las tecnologías de minería de textos cubren dos aspectos: el descubrimiento de conocimiento y la extracción de información.

3.2 Algunas aplicaciones típicas de MT

La empresa Statsoft [27] habla de que los textos no estructurados son muy comunes, de hecho representan la mayor información disponible en proyectos particulares de investigación. A continuación se describen brevemente algunas aplicaciones de la MT.

Análisis de encuestas con preguntas abiertas. Con ello, se pueden descubrir un cierto grupo de palabras que son comúnmente usadas por los encuestados, para describir pros y contras de cierto producto o servicio, para sugerir malinterpretaciones que son comunes y para encontrar elementos que en las preguntas estructuradas no se puede hacer.

Procesamiento automático de mensajes, correos electrónicos, etc. Con estas aplicaciones es posible filtrar automáticamente correo basura, basado en ciertos patrones que indican una fuente no reconocida. También es posible construir sistemas automáticos para clasificación de mensajes electrónicos y enrutamiento de los mismos a los departamentos o áreas mas apropiadas, mas aun es posible detectar palabras ofensivas o altisonantes, de modo que se puede responder al emisor un extrañamiento y solicitud de que se reenvíe sin ofensas.

Análisis de garantías, pólizas de seguro, entrevistas de diagnóstico, etc. Cuando alguien lleva su auto al servicio de mantenimiento, generalmente el que atiende escribe un reporte de los problemas detectados, este reporte puede convertirse a texto electrónico y conformarse como entrada a un sistema de MT que permita identificar grupo de problemas y quejas comunes a ciertos vehículos.

Investigación de los competidores mediante la revisión de sus sitios Web. El procesar automáticamente el contenido de un sitio o pagina web en un dominio particular se esta convirtiendo en una aplicación gran popularidad. Por este medio, es

posible derivar una lista de términos y documentos disponibles en el sitio y por ende determinar las características más importantes contenidas en el sitio que permitan descubrir las actividades de los competidores.

3.2.1 Grupo de trabajo 3. Aplicaciones de minería textual y manejo de conocimientos.

Según lo moderado por Bolasco del Grupo de Trabajo 3 de NEMIS (Network of Excellence in Text Mining and its applications in Statistics) [28], la habilidad para destilar información desde diferentes fuentes provee sustanciales ventajas competitivas para una compañía que desea tener éxito en las áreas de Administración de Conocimientos e Inteligencia de Negocios. La MT puede aplicarse en:

Análisis de mercados, en el cual un mercadólogo recoge estadísticas sobre la ocurrencia de palabras, frase o temas que serán útiles para estimar mercados, curvas de demanda y demográficas, revisión de detalles de productos desde catálogos en línea.

Manejo de relaciones con el cliente (CRM), mediante la minería de correos de entrada de las quejas y retroalimentación de clientes, aunque también se puede realizar en llamadas telefónicas, foros, listas de correo, cartas, encuestas de opinión.

Administración de recursos humanos, en la cual con la minería de reportes de la compañía y sus correspondientes actividades, condiciones y problemas, se pueden detectar a tiempo las variables que pueden indicar el estado del clima organizacional.

Análisis de bases de datos de patentes para detectar los mayores participantes, tendencias y oportunidades.

Vigía tecnológica para extraer eventos de negocios desde sitios web y crear una base de datos investigables para lo último de información de ciertos negocios.

Diseminación de información para organizar y resumir noticias y reportes de servicios de información personalizada.

3.2.2 Minería textual de Ciencia y Tecnología

De acuerdo con lo publicado en el GIT [29], es la aplicación de MT con el material técnico más altamente detallado. Sus principales componentes son: Extracción, Procesamiento e Integración. El procesamiento a su vez posee tres elementos: la bibliométrica, lingüística computacional y técnicas de agrupamiento.

Las aplicaciones principales en el área son: Lectura de documentos, identificación de infraestructura, identificación de temas y sus vínculos, descubrimiento de literatura y pronósticos de tecnología.

La bibliométrica es una técnica para capturar el número de veces que aparecen ciertas publicaciones, patentes o citas para medir e interpretar el avance de la ciencia. De ahí, puede ser usada como indicadores de innovación en cierta tecnología. El mapeo del flujo de innovación es una técnica para modelar las influencias o impulsores del desarrollo tecnológico en forma gráfica. Se usa para examinar la prospección de las áreas tecnológicas. La inteligencia tecnológica competitiva es usada para conocer las capacidades tecnológicas han desarrollado los competidores en cierta tecnología.

3.2.3 Minería textual en el campo de la mercadotecnia

Según Ravindran [30], los profesionales de inteligencia de mercados han estado usando diversas tecnologías para analizar numerosos productos, invertir en negocios, identificar necesidades de clientes, posicionar productos y otros. Herramientas analíticas basadas en tecnología TM se están popularizando para encontrar pepitas de bases de datos no estructuradas y para comunicar estrategias a los ejecutivos mediante la visualización de esos hallazgos.

El conocimiento de información de mercado está disponible en la Web a través de papeles blancos, publicaciones académicas relativas a mercados, periódicos especializados, artículos de noticias de mercados y opiniones del público que conducen a requerimientos de los clientes.

La MT ayuda a definir relaciones entre diferentes palabras clave mediante algunas técnicas como: agrupación de conceptos, indexado, asociacionismo, extracción de características, visualización de información y síntesis.

Estas técnicas describen conceptos y patrones a través de varias fuentes para apoyar a los profesionales de mercadotecnia a identificar información escondida que conduzca a aprovechar oportunidades de negocio.

El procedimiento aconsejado consiste en construir en primer lugar una biblioteca digital en el área de interés. En segundo término es necesario preprocesar los archivos resultantes mediante reducción de palabras, distinguiendo sinónimos y polisemias. Por último, se hace la interpretación de los textos minados para la detección de patrones.

3.2.4 Evaluación de aplicaciones de minería textual

De acuerdo a lo publicado por Second Moment [31]. Actualmente más empresas se agregan a otras que están aplicando tecnologías de MT, desde la mejora en la calidad de servicios al cliente hasta cuestiones de interpretación legal. Como toda tecnología nueva, existen problemas en su implementación, razón por la cual se presentan las siguientes características, las cuales han de ser evaluadas cuando se intente implementar aplicaciones de MT:

- Escalabilidad al crecer los datos
- Procesamiento desde múltiples fuentes

- Compatibilidad con sistemas existentes
- Capacidad de actualización y categorización automáticas
- Síntesis de grandes cantidades de textos
- Representación visual para una mejor comprensión
- Diversas técnicas de análisis textual
- Manejo de varios idiomas

La lista anterior puede ser un punto de partida para identificar y evaluar las herramientas más efectivas de acuerdo a las necesidades de la empresa.

3.3 Proyectos de minería textual para el tratamiento de noticias

3.3.1 Clasificación automática de textos de desastres naturales en México

Conforme a Téllez-Valero y otros [32], este proyecto tuvo como objetivo el desarrollar un sistema de clasificación automática de noticias de desastres naturales en México. Los resultados que se obtuvieron permiten clasificar una página web dentro de las categorías de huracán, inundación, sequía y no relevantes con una exactitud aproximada del 97%.

El proceso que se siguió fue:

1. Buscar y conformar el conjunto de documentos para el entrenamiento
2. Extraer del conjunto las características principales (conjunto de palabras)

Preprocesamiento (Limpieza de documentos)

Eliminar etiquetas HTML

Eliminar signos de puntuación

Quitar palabras insignificantes

Reducir las palabras a su raíz léxica

Indexado (Conformación de vectores)

Reducción del tamaño (dimensionalidad) de los vectores

3. Clasificar mediante los métodos vecinos más cercanos y clasificador simple de Bayes.
4. Evaluar los resultados y comparar con estándares mediante validación cruzada con 10 pliegues

Al final se determinó que el mejor camino para esta instancia fue aplicar la técnica de ganancia de información para reducir la dimensionalidad y además, aplicar el algoritmo simple de Bayes para llegar a la cifra de 97% de exactitud. Dentro de los trabajos futuros está el ampliar el número de clases que se manejan; construir un sistema adaptable a los ejemplos de prueba mal clasificados; continuar experimentando con otros métodos de representación de documentos y aplicar otros algoritmos de clasificación.

3.3.2 Un sistema para la clasificación de notas periodísticas

De acuerdo con García-Menier [33], este trabajo tuvo como propósito brindar una herramienta para la clasificación de noticias, debido a que es demasiada la información que se genera día a día y la capacidad para analizarla por el ser humano no es tan rápida.

Se pensó en un sistema que en base a ciertos criterios, clasificara las notas como negativas o positivas para una entidad específica.

Las técnicas actuales para la clasificación de textos son:

- Representar los documentos como un espacio vectorial y medir la distancia entre ellos.
- Utilizar una red neural.
- Utilizar técnicas estadísticas.

Estos métodos presentaban el problema de la excesiva dimensionalidad de los vectores, el cual se atacó usando la indexación por semántica latente. Existe otro método para reducir la dimensionalidad, conocido como *Clustering* Distribucional, que agrupa palabras con base a la distribución de la probabilidad que aparezca la palabra en el documento.

Existen Sistemas que toman frases completas y en base a esto realizan la clasificación de la nota ya sea como positiva o negativa.

Los pasos para la clasificación de la noticia son:

- Identificación de palabras clave o datos relevantes, para eliminar información innecesaria, para clasificar estas palabras se construyó un diccionario de manera manual.
- Agrupar palabras o frases sinónimas, para facilitar la calificación de frases equivalentes.

- Unidades semánticas (basadas en la metáfora de la persona con debilidad auditiva).

Las unidades semánticas son las que van calificando la nota, para hacer este proceso, se clasifican las palabras por separado de las unidades semánticas y en conjunto, es decir, dependiendo de las palabras que se encuentren. Ejemplo si se encuentra la palabra “muertes”, la calificación será mala, si en conjunto con otra palabra se tiene “disminuyen muertes” entonces cambiara a buena.

- Contar las notas positivas y negativas en el artículo, para poder calificar si la nota es positiva o negativa para la entidad.

Si el conteo de negativas es mayor que el doble de positivas, la nota se considera negativa, de lo contrario, la nota se considera positiva.

Este tipo de herramienta se puede usar para cualquier área que requiera algún tipo de clasificación de textos y no necesariamente periodísticos, podría ser de comentarios que se publiquen en foros de Internet.

3.3.3 Un sistema de descubrimiento de tendencias para minería de contenido de web dinámico.

En este artículo, los autores Méndez-Torreblanca y otros [34], describen un método completo para la minería de noticias en línea, el cual navega a través de estos sitios Web, extrayendo las noticias de ellos.

Básicamente el método propuesto es adaptado para la característica peculiar de las noticias. Aplica esquemas dinámicos para la extracción de reportes noticiosos y estrategias estadísticas independientes de dominio para la identificación de tópicos y análisis de tendencias.

El método propuesto, como cualquier otro método minero de la web, consiste de cuatro grandes fases: identificación del recurso; preprocesamiento, generalización y análisis, las cuales se muestran en la figura 3.2 y se describen enseguida.

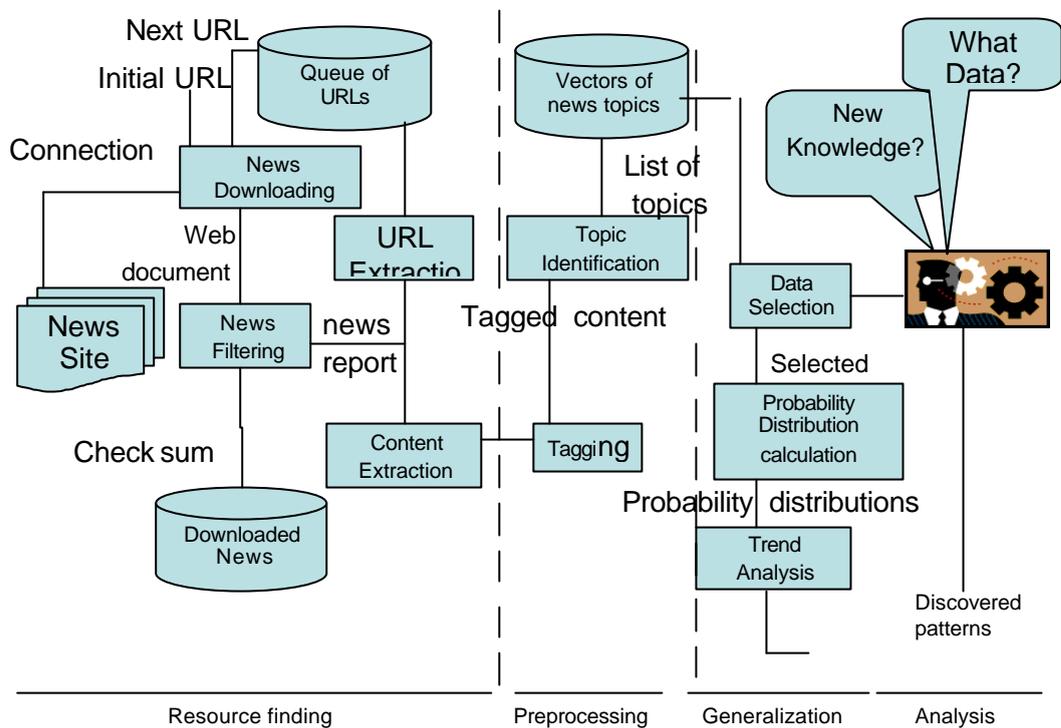


Figura 3.2 Método propuesto por Méndez-Torreblanca.

3.3.3.1 Identificación del recurso

Esta fase extrae todos los informes de las noticias de un sitio dado. Se lleva a cabo como un reptil que navega por un sitio de la red y continuamente extrae los informes de las noticias de él. Este reptil opera como sigue:

- Descarga la página del sitio web donde se encuentran las noticias
- Filtra la página descargada. Elimina la página si ya había sido descargada o pertenece a otro sitio. También verifica si la página es un informe de las noticias y no está anticuado, volviéndose entonces un documento de interés.
- Analiza el informe de noticias. Elimina información no pertinente como etiquetas y almacena el contenido del informe de las noticias para su proceso.

El proceso de identificación se activa periódicamente, ej. diariamente o dos veces en un día. Después de que este proceso se completa, los documentos obtenidos constituyen una foto instantánea de los eventos actuales y son preprocesados y almacenados. Después, el usuario puede seleccionar un periodo para el análisis.

3.3.3.2 Preprocesamiento.

Esta etapa transforma los reportes noticiosos entrantes en una representación estructurada. Se reduce el texto a una lista de tópicos, los cuales son vinculados a sustantivos y almacenados los más populares.

3.3.3.3 Generalización.

Se trata de descubrir tendencias interesantes entre los tópicos mediante dos actividades: la construcción de las distribuciones de los tópicos y el análisis de tendencias.

3.3.3.4 Análisis

En esta etapa el usuario interactúa con el sistema, estableciendo el período de trabajo y algunos parámetros de control. El usuario puede analizar entonces los patrones descubiertos y si no le son interesantes, puede repetir el proceso con otro período y parámetros hasta encontrar resultados satisfactorios.

Los resultados indican que el descubrimiento de tendencias en las noticias ayuda a interpretar los intereses de la sociedad y descubren información escondida acerca de la relación entre los eventos de la vida social.

3.3.4 Encontrando los principales temas en un documento escrito en Español.

Guzmán [35] ofrece el desarrollo del Algoritmo CLASITEX, el cual analiza un documento conteniendo lenguaje natural con la intención de reconocer los principales temas. Se desarrolló en el laboratorio de Lenguaje Natural y Procesamiento de Texto del Centro para la Investigación en Computación del Instituto Politécnico Nacional.

El análisis de textos en español permitirá que la computadora entienda los documentos escritos en lengua natural, por ejemplo, para resumirlos, encontrar tendencias, comparar dos documentos (con respecto a un tema dado) y contestar a preguntas no triviales.

El reporte inicia presentando la diferencia entre palabra y concepto, indicando que las palabras son ambiguas y los conceptos no. Por tanto, se busca trabajar con los conceptos que representan las palabras, mediante el análisis semántico que trata con los significados.

Los conceptos forman familias o jerarquías, en las cuales la relación más usada es el subconjunto, donde los hijos denotan subconjuntos o especializaciones o conceptos más específicos que el padre. El trabajo consiste en formar un árbol o una parte del mismo para conformar un número finito de conceptos. El árbol puede organizar el conocimiento común, el cual poseen aquellos que no manejan una especialización y está conformado por unos diez millones de conceptos.

El árbol del sentido común tiene, además de "subconjunto", otras relaciones (tales como "parte de", "formado por" y otras) entre dos conceptos. Las relaciones son también parte del árbol de conocimiento.

A partir del árbol, se lee el documento y se reemplazan las palabras por sus respectivos conceptos. Conforme se encuentran los conceptos se incrementan sus contadores y los de aquellos que son sus padres hasta el concepto de mayor nivel. Al final del análisis, los conceptos del árbol de mayor frecuencia serán los tópicos principales del documento.

Los problemas que presenta el método propuesto es que el árbol tiene que construirse manualmente y además es necesario eliminar la ambigüedad de la palabra, pues una misma puede referirse a varios conceptos. Sin embargo, luego de varias pruebas se encontró que saltarse el proceso de desambiguación no afectaba la precisión de los resultados, pues al final imperaban los conceptos más populares.

El verdadero trabajo estriba en el análisis de conceptos representados por más de una palabra, pues el algoritmo tiene que trabajar con las palabras actual, previa y siguiente para verificar si un grupo representan un concepto (ej. Benito Juárez representa el concepto presidente).

Los resultados encontrados indican que conforme el documento por revisar es más corto es difícil precisar los principales tópicos. Se han procesado textos provenientes de revistas y periódicos genéricos, lo cual requiere de un árbol menos complejo. Para tratar textos más especializados se requiere enriquecer en gran medida el árbol conceptual.

3.3.5 Nominalización en títulos: una manera de extraer detalles de documentos.

De acuerdo con Montes y Gómez y López [36], este trabajo habla de que los títulos de los documentos son lo primero que el lector lee acerca de su contenido. A partir de ello, la nominalización en títulos proporciona un significado para identificar un nivel de contenido específico como es sugerido por el verbo del cual se derivan.

Comúnmente los documentos científicos se expresan como un conjunto de palabras clave, lo cual no es muy preciso pues deja fuera algunos aspectos importantes del documento. Con la explosión de la información por medios electrónicos, como el Internet, se requiere de mejores métodos de búsqueda y clasificación de documentos.

Se propone un nuevo sistema de búsqueda que utiliza otros elementos además de los términos para indexar los documentos y obtener mejor información. El sistema es un método de representación de dos niveles que permite incluir los detalles de contenido con su estructura, es decir como se relacionan los conceptos. Se toman los verbos que se encuentran en el título y se explora como sus nominalizaciones son usadas, para evaluar características del contenido.

Se llevó a cabo un análisis manual de 4663 documentos de las colecciones CACM y CISI. Un listado de verbos fue usado para buscar los sustantivos derivados de ellos. La nominalización es a menudo usada para expresar información no relevante, por ejemplo en “Una introducción a una división de datos independiente de la máquina” se puede deducir que el documento introduce a “división de datos”, en un método normal solamente se guardaría “división de datos” y no el nivel básico al que se refiere con “introducción”.

Se encontró que no todas las nominalizaciones en títulos expresaron la clase de detalles de interés, pues solamente aquellas nominalizaciones del documento que provinieron de los verbos que articulan las intenciones del documento produjeron resultados correctos. Los títulos analizados indican intenciones de diversas actividades. Éstas generalmente se relacionan el contenido específico de los documentos, como el objetivo, punto de vista entre otras.

El proceso consistió en el pre-procesamiento para aislar el título, el etiquetado del título, el filtrado, la división del título en sus partes, la generación del grafo conceptual y la recuperación de datos implícitos.

Los resultados obtenidos con el análisis de 4663 documentos fue que se encontraron títulos con nominalización en un 14.62% (682), mientras que presente participio hallado en un 5.16% (241) dando un total de documentos útiles del 19.79% (923).

Se encontró evidencia de que la nominalización es una manera confiable de acceder a más información acerca del contenido del documento, mediante un medio de relacionar las expresiones en títulos que puede ser también usada para acceder a información similar. La representación estructurada de la información extraída de los títulos de los documentos proporciona detalles específicos de su contenido. Mediante una representación de dos niveles de dimensión en los cuales un documento puede ser juzgado, sirviendo a un amplio rango de interés.

3.3.6 Minería de Noticias: Tendencias, Asociaciones y Desviaciones

M. Montes-y-Gómez et. Al [37], reportan haber desarrollado un método para la minería de noticias. Éste busca analizar las colecciones de noticias aparecidas en periódicos y otros medios masivos. El interés surge por la fuerte relación que existe entre los intereses de la sociedad y los tópicos de las noticias. Las noticias son muy diversas de cambio constante y se influyen entre sí.

El método utiliza una representación estadística sencilla de los reportes noticiosos para la detección de tendencias, desviaciones y asociación efímera entre temas. A continuación se describen brevemente cada una de estas representaciones.

Análisis de Tendencias. El objetivo principal del análisis de tendencias es estudiar el comportamiento de los intereses de la sociedad, determinando si ellos cambian o permanecen estables de un periodo a otro. Para realizar el método se utilizan dos colecciones, por ejemplo: dos nuevas fuentes, los periódicos de dos sitios geográficos, etc. Se divide en dos tareas principales: Descubrimiento general de tendencias e identificación de temas o factores que contribuyen a esta tendencia.

Descubrimiento de asociación efímera. Un fenómeno usual en noticias es la influencia de un tema de interés máximo en las noticias, es decir, un tema a corto plazo de una sola vez de frecuencia máxima, sobre los otros asuntos de las noticias. Esta influencia se demuestra en dos diversas formas: el tema máximo induce a algunos asuntos que emerjan o lleguen a ser importantes junto con él, y a los otros que se olvidarán momentáneamente. Esta clase de influencias (relaciones de tiempo) es lo que se determina como asociaciones efímeras. Una asociación efímera se puede ver como relación directa o inversa entre las distribuciones de la probabilidad de los asuntos dados sobre una duración fija.

Detección de desviación. La detección de desviaciones en una enorme cantidad de datos es una tarea importante, pero difícil. Tiene como objetivo descubrir elementos irregulares en una gran cantidad de datos.

Resultados Experimentales. **Análisis de Tendencias.** Se consideraron noticias económicas del Universal en la última semana de Enero y la primera semana de

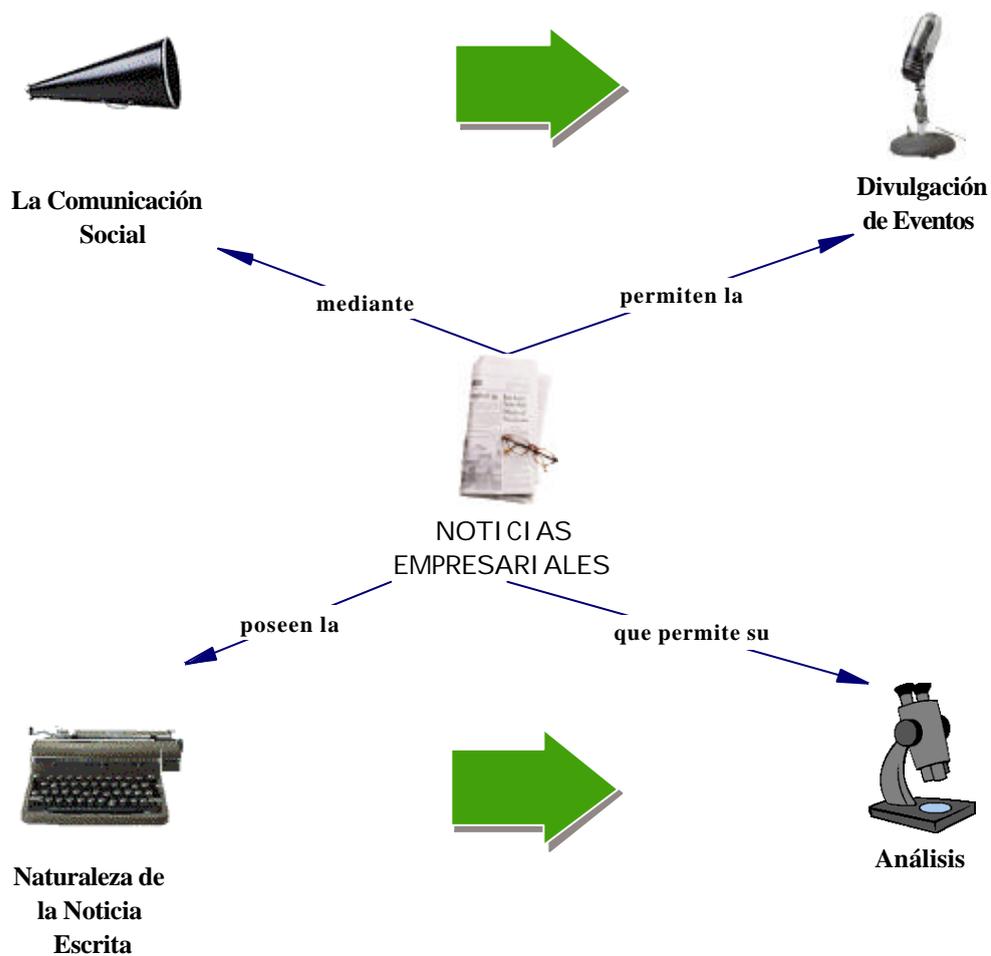
Febrero de 1999. Hubo 47 temas en esas dos semanas, de las cuales de un filtrado y fusión de estos, se obtuvieron solo 15 temas. Con sus distribuciones de probabilidad se detectó un cambio global de tendencia entre estos dos periodos. Descubrimiento de Asociaciones Efímeras. Se consideró la noticia nacional de la visita del Papa Juan Pablo II a la ciudad de México, como tema principal, en el Universal por diez días. Se detectaron dos asociaciones efímeras: Una asociación entre el tema principal y la Virgen de Guadalupe. Por otro lado hay una asociación inversa entre el tema principal y el tema de Raúl Salinas de Gortari. Esto indica que el tema de la Virgen se vuelve importante, mientras que la otra asociación pudo haber causado el olvido de la sentencia de Raúl Salinas.

Los informes de noticias son una fuente importante de información acerca de la sociedad. Su análisis permite entender mejor sus intereses actuales y para medir la importancia social de muchos eventos. En este método solo se utilizan las medidas estadísticas en el procedimiento de descubrimiento, tales como el promedio, la desviación estándar y coeficiente de correlación.

Capítulo IV.

La naturaleza de las noticias

Este capítulo trata sobre la importancia de la comunicación social, las características de la misma, en las instituciones del sector elegido, las implicaciones a considerar en la divulgación de sus eventos, la estructura que guardan los textos que reportan hechos noticiosos y los trabajos de análisis que se han realizado con la intención de convertir las noticias de su forma no estructurada a expresiones formales que permitan un manejo más fácil.



4.1 LA COMUNICACIÓN SOCIAL Y SUS FUNCIONES

De acuerdo con MONCADA [38], la comunicación social es un campo de estudio que explora principalmente las áreas de la información que puede ser percibida, transmitida y entendida, así como el impacto que puede tener en la sociedad.

Estas son algunas de las principales actividades de la Comunicación Social (Nuevo León 2003):

- Establecer acuerdos entre los altos mandos de la organización y el despacho de los asuntos inherentes de la unidad administrativa a su cargo y responsabilidad, e informarle oportunamente sobre los mismos.
- Dirigir el enlace y la comunicación de la Corporación con otras organizaciones tanto gubernamentales como particulares, así como coordinar el trabajo con otras áreas de comunicación social.
- Dirigir y coordinar la realización de entrevistas, conferencias, boletines y, en general, las actividades de comunicación social de los actos o acontecimientos de la Corporación para su difusión en medios de comunicación, así como integrar las bases de datos de información de noticias en materia de desarrollo cultural.
- Emitir información a través de los diferentes medios de comunicación y relaciones públicas, sobre el desarrollo de los programas y actividades de la Corporación.
- Procesar y coordinar las acciones de difusión y divulgación de material informativo de los diversos sectores en medios masivos alternativos, así como

proporcionar al público solicitante información y orientación sobre las actividades, trámites y programas de la Corporación.

Conforme lo indica PIÑUEL [39], se entiende por Sistema de Comunicación Social a un sistema peculiar de comunicación en el que el intercambio de expresiones entre actores colectivos (los titulares de los Medios como emisores, y los públicos o audiencias como receptoras) se somete a los procesos de producción y consumo de un servicio convertido en mercancía y que consiste en facilitar, por el intercambio de datos codificados y decodificados por los actores/agentes sociales, la reproducción de conocimientos a propósito del acontecer social y material, que es un objeto de referencia frente al cual los miembros de la colectividad social necesitan ajustar sus conductas como agentes y como sujetos con aspiraciones sociales y materiales.

4.1.1 La comunicación social en las Universidades Privadas

Los departamentos de comunicación social en las Universidades Privadas en México, se encargan de proporcionar la oferta académica de estas a través de medios publicitarios y estrategias de Venta Directa (promoción directa en colegios, estantes en exposiciones y en plazas comerciales) dirigidas al público, tanto interesados (aspirantes) como a los padres de familia.

Además dan difusión a las actividades de las diversas áreas que puedan ser del interés de la sociedad y también se encargan de unificar la imagen institucional.

La comunicación social extrainstitucional la cual difunde entre la sociedad, el desarrollo académico, científico, tecnológico y artístico de la institución, como depositaria de los valores culturales y artísticos, de la comunidad en la cual se desenvuelve.

4.1.2 La divulgación de eventos.

De acuerdo con Rodríguez [40], los nuevos productos y servicios crean nuevos mercados y la publicidad se convierte en una de las principales herramientas para darlos a conocer y provocar su demanda. Por lo anterior tenemos que una empresa, debe publicar sus novedades, entre las que pudieran ser, nuevos productos o servicios.

El mismo autor refiere que un texto informativo es un tipo de texto a través del cual el emisor da a conocer a su receptor algún hecho, situación o circunstancia. Remarcando los beneficios para el usuario tanto personales como emocionales. Esto depende de que es lo que se desea comunicar:

- Producto
 - Propósito del producto
 - Atributos
 - Calidad y Valor
 - Usuarios
 - Usos
 - País de origen
- Organización
 - Atributos de la organización
 - Fortalezas y oportunidades.

Rodrigo [41] menciona que el establecimiento de un acontecimiento público depende de tres actores: los promotores de noticias (las fuentes); los recolectores de noticias (los periodistas) y los consumidores de noticias (la audiencia). Atendiendo a las expectativas de los lectores y de la sociedad en su conjunto, y a la propia naturaleza de los medios de comunicación de masas en cuanto a mediadores sociales, el proceso noticioso tiene dos objetivos fundamentales:

Responder a las necesidades informativas del público, entendidas éstas como “el conjunto de informaciones que precisa un ciudadano para desenvolverse con autonomía de juicio y libertad de acción en la sociedad en que vive”. Fontcuberta [42].

Satisfacer los requerimientos de la colectividad, la cual requiere que los medios proporcionen cada día aquellos temas y hechos “que generan conexiones y refuerzan el sentido comunitario” Rozas [43].

El momento más adecuado para divulgar una noticia, es cuando se da alguna de las situaciones de la tabla 4.1:

Lanzamiento de un nuevo producto o servicio.	Mejoras técnicas, de calidad o de organización en su servicio.
Participación en ferias, congresos o conferencias.	Consecución de nuevos clientes de reconocido prestigio.
Acciones ecológicas o encaminadas a mejorar la seguridad.	Obtención de premios o reconocimientos.
Expansión a nuevos mercados o líneas de negocio.	Ampliaciones de personal
Firma convenios de colaboración.	Informe sobre resultados económicos.
Convocatorias de prensa.	Participación u organización en actos benéficos.
Participación u organización en actos culturales.	Participación u organización en actos deportivos.
Participación u organización en actos sociales.	Obtención de premios, nominaciones y reconocimientos.
Donaciones.	Patrocinios.
Evolución positiva de su cifra de ventas.	Alianzas empresariales.
Apertura de nuevas filiales.	Compra de empresas de la competencia.
Inversiones en nuevos negocios.	Obtención de certificados de calidad (ISO, AENOR, etc.)
Aniversarios de productos o servicios.	Aniversarios de la creación de la propia empresa.
Descuentos y rebajas.	Desmentido de rumores y calumnias.
Comentario de sentencias judiciales.	Comentarios ante problemas coyunturales de su sector.
Realización de jornadas de puertas abiertas.	Presentaciones comerciales.
Participación en seminarios o conferencias.	Declaraciones destacadas de directivos.
Contrataciones y ceses de directivos.	Resultados sobre estudios de mercado.
Presentación de informes comerciales.	Realización de estudios de impacto social.
Resultados de investigaciones científicas.	Consecución de mejoras para la sociedad.
Nuevo nombre o imagen corporativa.	Ampliaciones de capital.
Anuncio de futura junta de accionistas.	Resultados de juntas de accionistas.
Previsiones de facturación.	Publicación de anuncios oficiales.
Publicación de resoluciones judiciales.	Consecución de un récord de cualquier tipo.
Implementación de un programa de innovación.	Presentación de una campaña de publicidad o marketing.

Tabla 4.1. Situaciones adecuadas para divulgar una noticia.

Para Romañach [44], la comunicación social estructurada, hace referencia a la transmisión a la sociedad completa de algún tipo de experiencia, idea, proyecto o documento. En este sentido, la herramienta más útil y efectiva es la Web (World Wide Web).

4.2 LA NATURALEZA DE LAS NOTICIAS ESCRITAS

Shabb [45] señala, en la “Teoría de la Comunicación, el mensaje resulta de la acumulación armónica de tres elementos: el código (o sistema de signos utilizado), la forma expresiva y el contenido.”

Además indica que cada uno de los canales por los cuales hoy día se difunden los mensajes periodísticos tiene su propio sistema de signos, su propio código de señales para expresar ideas, para dar forma inteligible a los contenidos, esto es: su lenguaje. En éste lenguaje del periodismo escrito, se reconoce la existencia de varios géneros periodísticos perfectamente diferenciables entre sí:

- El género informativo
- El género interpretativo
- El género editorializante.

Cada uno de los géneros -el informativo, el interpretativo y el de opinión- tiene su propia técnica de trabajo y su propia estructura o formato.

En el género informativo, el objetivo es informar y captar al lector por el camino de la comunicación de unas noticias y la hábil exposición de ideas, que se consigue por diferentes caminos y que a su vez dan origen a diferentes modalidades de estilo periodístico

Los géneros puros pertenecientes al género informativo son la Noticia y la Crónica.

La Crónica es el relato extenso de la noticia, abundando en causas, detalles, incluyendo descripciones temporales o ambientales, retratos, narración, reproducción de discursos directos, con la posibilidad de incluir observaciones subjetivas. Generalmente redactado por un periodista testigo de los hechos, en lenguaje informativo.

La Noticia consiste en la escueta enumeración de los datos de un acontecer a producirse y producido. También se puede considerar a la noticia como el género periodístico por excelencia que da cuenta, de un modo sucinto pero completo, de un hecho actual o actualizado, digno de ser conocido y divulgado, y de innegable repercusión humana. Martínez [46] abunda, pues señala que: "La Noticia es un hecho verdadero, inédito o actual, de interés general, que se comunica a un público que pueda considerarse masivo, una vez que ha sido recogido, interpretado y valorado por los sujetos promotores que controlan el medio utilizado para la difusión".

La estructura de los textos es un factor fundamental en la facilitación de la comprensión del texto por parte de los lectores. Una noticia en su estructura contiene:

TITULAR Y ENCABEZAMIENTO

Titular: puede contar con el título principal, sobretitular o volanta (arriba) subtitular o bajada (abajo).

Encabezamiento :

Consiste en el:

-Copete

-Resumen en primera oración: lead o cabeza

SITUACION: Episodio y Antecedentes

-Acontecimientos principales

-Acontecimientos consecuentes

-Antecedentes: información del contexto o sucesos previos

-Historia: sección de un texto que remite a situaciones no recientes

SUCESOS PREVIOS O CONTEXTO: para recordar al lector hechos ocurridos previamente.

Comentario: reacciones verbales y conclusiones

REACCIONES VERBALES U ORALES: citas directas e indirectas que legitiman el discurso periodístico.

CONCLUSIONES: Evaluación y opiniones

- Evaluación: permite opinar sobre acontecimientos principales.
- Expectativas: formulan consecuencias futuras de los acontecimientos principales.

Estas categorías periodísticas mencionadas definen todas las formas posibles del discurso periodístico informativo. Sólo el titular, el encabezamiento y los acontecimientos principales deben hallarse. Los antecedentes, reacciones verbales y comentarios son opcionales.

4.2.1 El lead

En castellano debiera traducirse este vocablo por arranque, entrada, comienzo de un texto informativo. En el caso de la noticia, el párrafo inicial busca condensar sinópticamente toda la noticia en aquellos datos esenciales para una cabal comprensión de la misma. Para escribir un correcto lead informativo hay que tratar de responder, y por orden de importancia relativa, las preguntas que están encabezadas por la letra W las cuales constituyen el signo gráfico inicial de los siguientes pronombres y adverbios en el idioma inglés.

Who - quien

What - que

When - cuando

Where - donde

Why - por que

How - como

4.2.2 La divulgación de la noticia.

Cuando difunde información a través de los medios de comunicación hay que recordar que la imagen que transmita en la noticia será la base de la imagen que los lectores finales tendrán de quien se habla en ella. La calidad del contenido y su veracidad, son básicos para captar la atención y la confianza del lector.

De acuerdo con Castresana [47], se tiene una serie de cuestionamientos para que la noticia sea considerada completa.

¿Quién? ¿Quién realiza lo que se explica en la noticia?, la propia empresa, sus empleados, un departamento determinado, etc.

¿Qué? ¿Qué sucede o cuál es el tema principal que quiere destacar?, el lanzamiento de un nuevo producto o servicio, la apertura de una nueva filial, una fusión, la presentación de resultados, la participación en una feria o congreso, etc.

¿Dónde? ¿Dónde se localiza el qué de la noticia?, en la empresa, en un congreso, en una calle determinada, etc.

¿Cuándo? ¿Cuándo ha sucedido, sucede o sucederá el qué de la noticia?, ayer, hoy, mañana, dentro de una semana, en una fecha concreta, durante un período de tiempo, etc.

¿A quién? ¿A quién se dirige el qué de la noticia?, a los potenciales clientes, a los clientes actuales, a los ciudadanos en general, a otras empresas o instituciones, etc.

¿Cómo? ¿Cómo sucede el qué de la noticia?, a través de una conferencia, de una remodelación, de una puesta al día, en una rueda de prensa, en persona, etc.

¿Por qué? ¿Por qué sucede el qué de la noticia?, por petición de los clientes, a partir de un estudio de mercado, etc.

¿Objetivo? ¿Qué se pretende conseguir con el qué de la noticia?, obtener nuevos clientes, ampliar la cuota de mercado, abrir nuevos mercados, mejorar

la rentabilidad, potenciar la imagen externa, para alinearse con el mercado o la legislación, para innovar, etc.

Singularidad. ¿Qué tiene de singular o exclusivo? ¿Es único? ¿Qué le hace diferente al resto?

Precedente. ¿Es el primero en el Mundo, en el continente, en el país, en su estado o en su ciudad?

Personalización. ¿Se han incluido declaraciones de algún cliente o afectado que pueda darle valor añadido a la noticia demostrando lo que cuenta?

Independencia ¿Se incluye, la declaración de algún experto independiente reconocido en la materia tratada? Esto ayuda a ofrecer una imagen de credibilidad.

Curiosidad ¿Tiene un aspecto que despierta el morbo o la curiosidad?

Provocación ¿Es provocativo o rompe con las normas generalmente establecidas?

Saludable ¿Tiene repercusiones positivas para la salud?

Mejor ¿Es más seguro, más barato, más confortable, más atractivo, más rápido, etc.?

Ecología ¿Tiene repercusiones positivas para el medio ambiente?

Solidaridad ¿Tiene alguna implicación solidaria?

Conflictividad ¿Existe algún conflicto a favor?

Reconocimientos ¿Ha conseguido algún premio o nominación?

Regalos ¿Existe alguna promoción o regalo (siempre se publica)?

Cuantificación ¿Cuáles son los datos numéricos objetivos sobre lo que pasa?

Posicionamiento ¿Cómo se posiciona respecto a la competencia o respecto a lo que se hace normalmente?

Temporalidad ¿Durante cuánto tiempo sucedió, sucede o sucederá el qué de la noticia?

4.3 ANÁLISIS DE HECHOS NOTICIOSOS

Como lo menciona Moriello [48]: ‘Con la tecnología informática, todos los procesos inteligentes del ser humano se vieron, de alguna manera, afectados. Hoy en día, las máquinas sustituyen al hombre en ciertas actividades mentales específicas: algunos aspectos de la inteligencia humana como la memoria, la lógica, la toma de decisiones, no sólo pueden imitarse artificialmente sino que son más rápidas, poderosas y confiables. Sin embargo, no se trata, en realidad, de construir máquinas que compitan con el ser humano, sino que sean capaces de realizar tareas de cierto nivel intelectual para ayudarlo. Las máquinas liberan al hombre para que éste enfoque su energía hacia las actividades creativas, intuitivas, originales y no cotidianas.

También indica que “a pesar de que puede acceder a muchos más datos, la capacidad de procesamiento del hombre sigue siendo limitada y por esta causa actúa como factor paralizante. El cerebro humano tiene una ínfima capacidad de almacenamiento frente a la vorágine documental del mundo actual. Es completamente imposible para un ser humano mantenerse informado de todo lo que sucede en su área de actividad. El hombre no puede digerir tanta información. Es un fenómeno que está cambiando al mundo. Mucha de la información es importante, y uno tampoco puede seleccionar lo que es importante”

Del Valle [49] propone el análisis de las noticias, usando la gramática de casos como modo de lectura. En este análisis se detectan los siguientes elementos:

Acción. Debe ser detectada mediante abstracción de la macroproposición o significación profunda y última de la noticia. ¿Qué ocurre? ¿Cuál es el acontecimiento más genérico? Su rol es servir de vértice a todos los demás casos conceptuales. Todo texto gira en torno a una acción y, como regla general, todas las acciones explicitadas en un texto son subsumibles en otra más amplia. Si no lo fueran podrían tratarse como microtextos independientes.

Agente. Categoría que abarca seres vivos, humanos, animales o vegetales, microorganismos, instituciones como representaciones o colectivos humanos y, eventualmente, personificaciones y entes espirituales que desarrollen o ejecuten una acción.

Paciente. Categoría que cubre los mismos elementos que la anterior siempre que estos tengan, en el texto, el comportamiento de paciente, es decir, quien sufre concreta y directamente las acciones (y no sus consecuencias).

Objeto. En realidad es el mismo caso Agente referido a cosas o entidades no vivas que directamente desencadenen acciones.

Objeto pasivo. Siempre que exista el caso Objeto como categoría independiente del Agente, denominaremos a las cosas o entidades materiales (también inmateriales, en el caso de declaraciones) que sufren directamente las acciones de Agentes u Objetos y, en general, a lo realizado transitivamente, Objeto pasivo. Lo hecho, lo vendido, lo acordado, lo construido, lo dicho o declarado.

Instrumento. Herramienta o medio a través del cual se ejecuta la acción. Son objetos utilizados con función instrumental.

Modo. Procedimiento general o especializado y voluntario, no implícito en la acción utilizado en la realización de la acción.

Causa. Pueden ser causativos: 1) Los motivos inmediatos, 2) los antecedentes y 3) los contextos generales en los que la acción se inserta, siempre que estén explícitos.

Finalidad. Objetivos o intencionalidad de la acción, explícitos en el texto pero no obvios ni realizados.

Producto. Efectos, reacciones, consecuencias o resultados reales y palpables de la acción.

Lugar. Noción espacial o contexto geográfico en el que tiene lugar la acción.

Tiempo. Noción temporal o contexto cronológico si fuera necesario, en el que tiene lugar la acción.

Situación. Caso secundario pues depende de la existencia de Agentes u Objetos. La situación es involuntaria, atributiva y afecta física o psíquicamente a los actuantes. ¿En qué situación se encuentra el Agente/Paciente/Objeto/Objeto pasivo?

Asociativo. Caso secundario referido a acompañantes o elementos subsidiarios, nunca protagonistas directos, de los agentes y los objetos. ¿Con quién/qué se encuentra el Agente/Paciente/Objeto/Objeto pasivo?

4.3.1 Análisis lógico de noticias.

El Dr. Colle [50] presentó un trabajo en el cual daba cuenta de un procedimiento de análisis lógico-semánticos de hechos noticiosos, inspirados en estructuras del conocimiento.

Este trabajo pretende desarrollar y aplicar un procedimiento en el campo del análisis y registro electrónico de noticias, que pudiera significar un importante avance en el futuro por cuanto permitirá transformar las actuales Bases de datos en Bases de conocimiento, cuyo funcionamiento y aprovechamiento posible, es muy superior.

Así, el autor piensa que a futuro, será posible pensar que una base de conocimiento alimentada con notas periodísticas podría estar en condiciones de hacer una gran parte del trabajo periodístico de interpretación.

Para llevar a cabo el procedimiento es necesario asentar algunas definiciones básicas de los elementos propios de este dominio.

Formalmente se inicia con un análisis básico que consiste en determinar si el hecho noticioso es tangible o intangible; luego se pasa a determinar cada uno de los elementos básicos para continuar luego con el análisis lógico-semántico del hecho noticioso; para finalmente, concluir con el registro en la base de conocimientos de los elementos relevantes que podrá servir para futuras consultas.

Para dar una idea cabal del procedimiento, se le puede revisar en el mapa conceptual de la figura 4.1 y para mayor abundamiento se presenta un ejemplo del mismo.

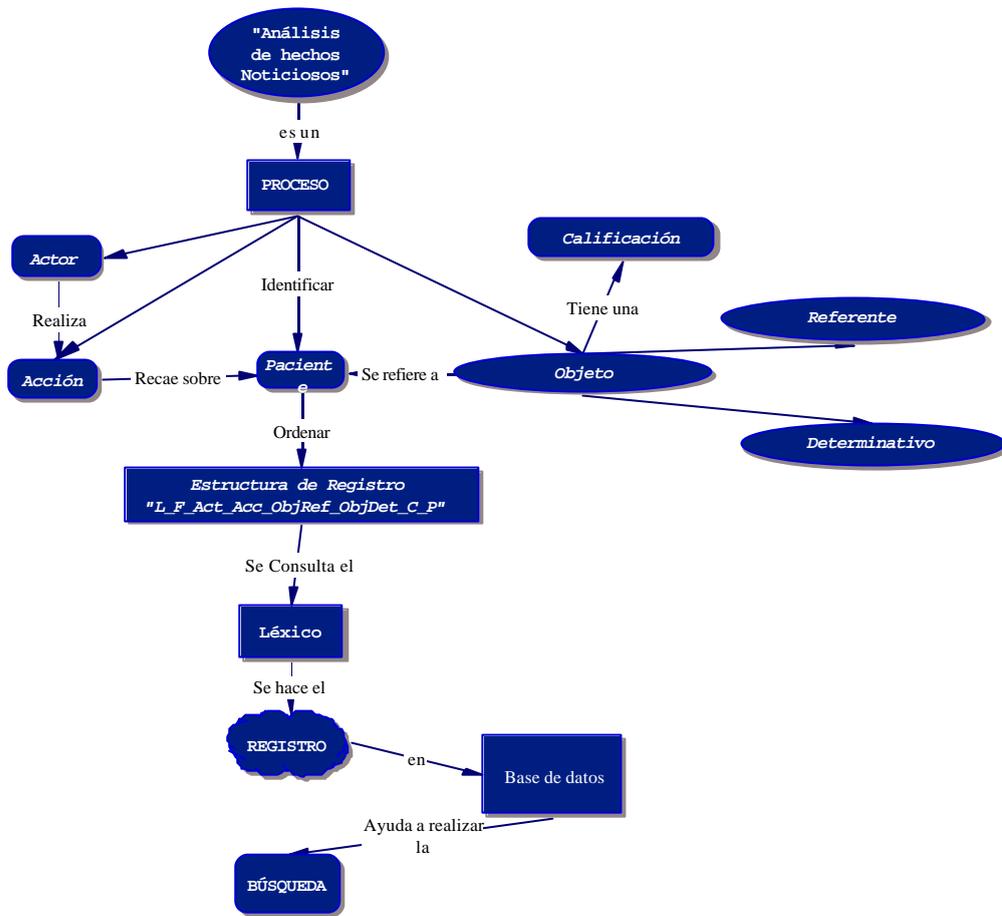


Figura 4.1 Mapa mental del procedimiento propuesto por el Dr. Colle.

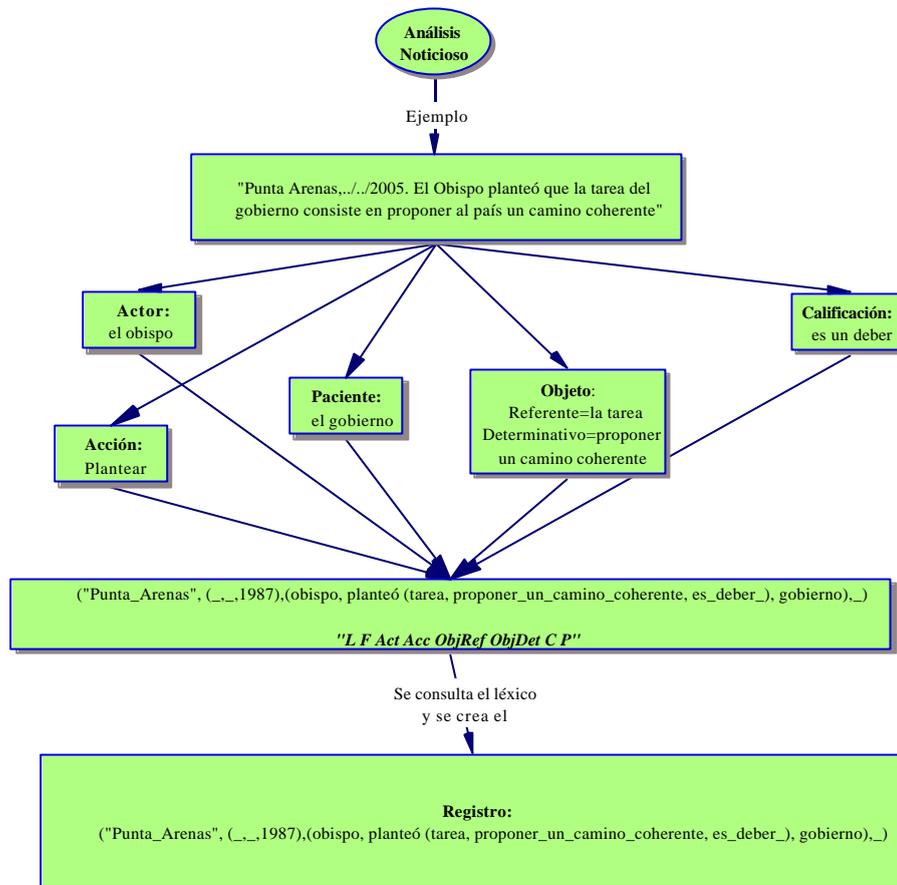


Figura 4.2. Ejemplo del procedimiento del Dr. Colle.

Capítulo V.

Propuesta de tratamiento automatizado de noticias empresariales

En este capítulo, a partir de la presentación de las características de las noticias empresariales publicadas en la Web, se describen los criterios que se tomaron en cuenta, para elegir el tipo y muestra de empresas por analizar. Además, se muestran las distintas corrientes seguidas en la transformación de textos a una representación semiestructurada y el criterio para la selección del usado en este trabajo. Se documentan los puntos sobre los cuales se fundamenta la manera de procesar las noticias. De todos los elementos que conforman un hecho noticioso, se definen adicionalmente, los criterios de selección de los elementos que han de procesarse. Se presentan los resultados preliminares de procesar los titulares de un día de publicaciones para determinar del total, la proporción de noticias elegibles. Previo a la presentación del procedimiento, se documentan las bases mínimas sobre las que trabaja la expresión humana en el idioma español. Luego, se documentan las etapas del procesamiento automatizado de las noticias y se exponen los ejemplos necesarios para una mejor comprensión. A partir del registro de los elementos detectados, se procede a mostrar las preguntas que, desde la inteligencia competitiva, se pueden formular al sistema computacional planteado y la manera como se responde a dichos cuestionamientos.

5.1 LAS NOTICIAS EMPRESARIALES EN LA WEB

Continuamente la empresa está publicando reseñas sobre las actividades que realiza, para fines publicitarios y de relaciones públicas. Están orientadas hacia sus diferentes públicos (clientes, proveedores, sociedad civil, etc.) pasados, actuales y potenciales, para promover una mejor imagen empresarial y posicionar su oferta en la mente del consumidor.

Para informar y ser informada, la empresa utiliza con mayor frecuencia la Web, por las ventajas económicas y de alcance que representa. Por su naturaleza, la comunicación multimedia de las noticias se hace más patente en las páginas Web. Las noticias publicadas en los portales, están vinculadas con el comportamiento que desean informar, permanecen publicadas más de un día y generalmente son de tamaño reducido.

La publicación es una tarea encomendada generalmente a los departamentos de mercadotecnia, publicidad, relaciones públicas, comunicación social y eventualmente al administrador del sitio (*Webmaster*).

Una correcta redacción de hechos noticiosos, requiere que la esencia del comunicado se presente en el título o por lo menos en el encabezado (*lead*).

De una revisión preliminar a las noticias de las empresas seleccionadas, se encontró que utilizan alguno de los siguientes formatos:

- Una lista de titulares de los eventos
- Una descripción breve y completa de los hechos
- Una lista de títulos con una imagen asociada a textos completos
- Un listado de títulos de eventos con enlace hacia páginas que informan a detalle
- Una relación de encabezado y descripción completa del evento
- Una lista de noticias con encabezado, una descripción parcial y un enlace a otra página complementaria.

Enseguida se muestran las imágenes correspondientes a cada uno de ellos.

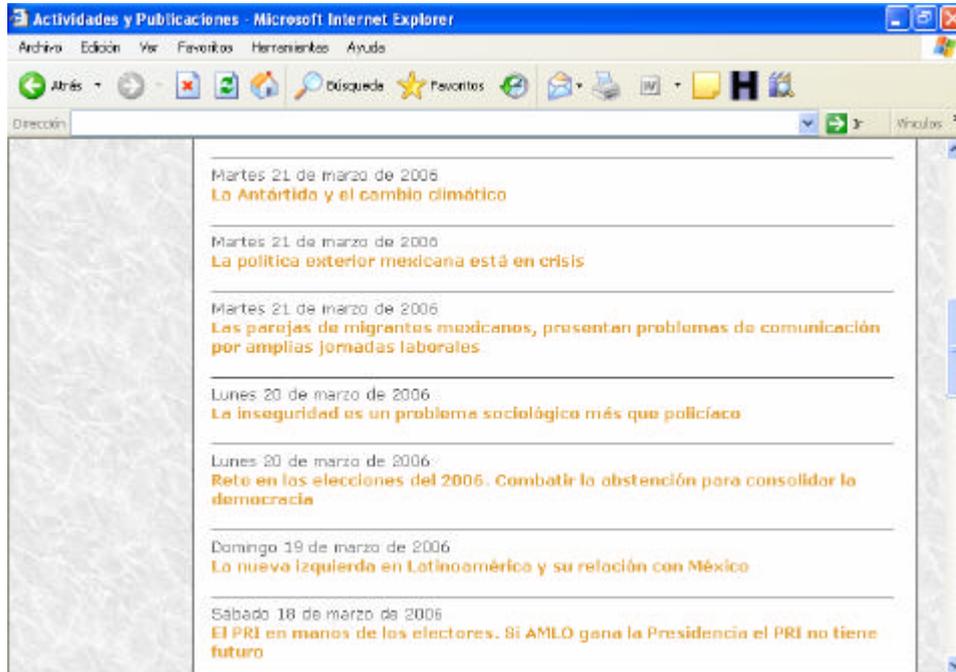


Figura 5.1. Una lista de titulares de los eventos

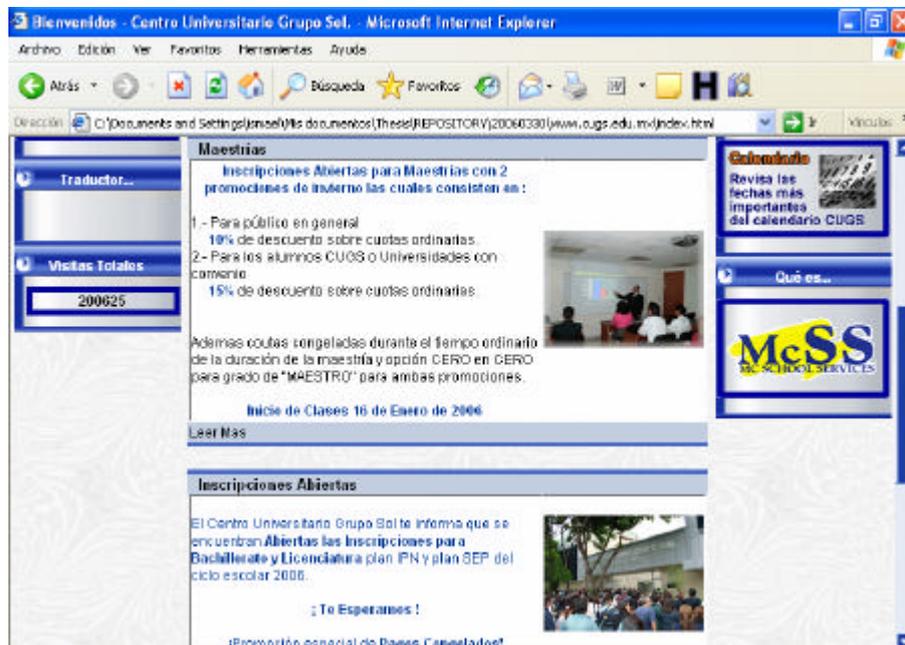


Figura 5.2. Una descripción breve y completa de los hechos

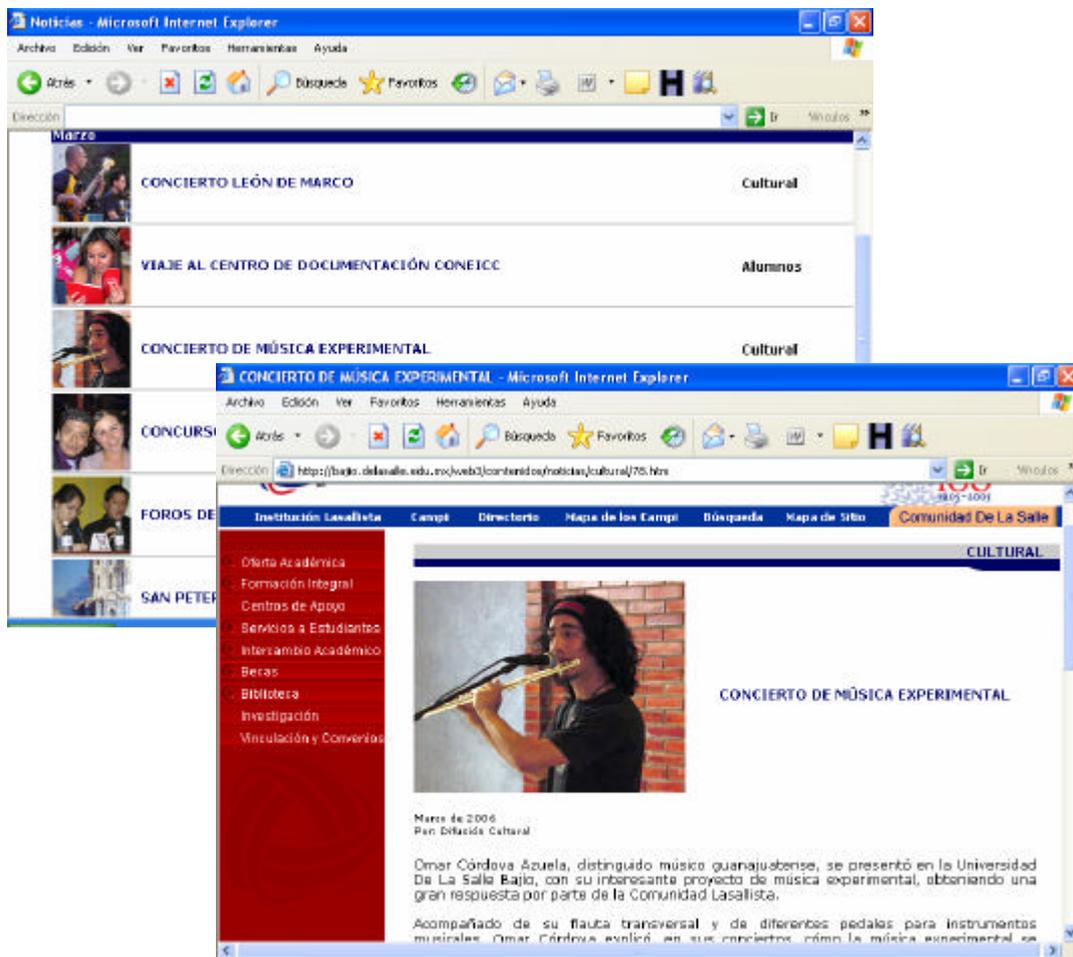


Figura 5.3. Una lista de títulos con una imagen asociada a un texto completo

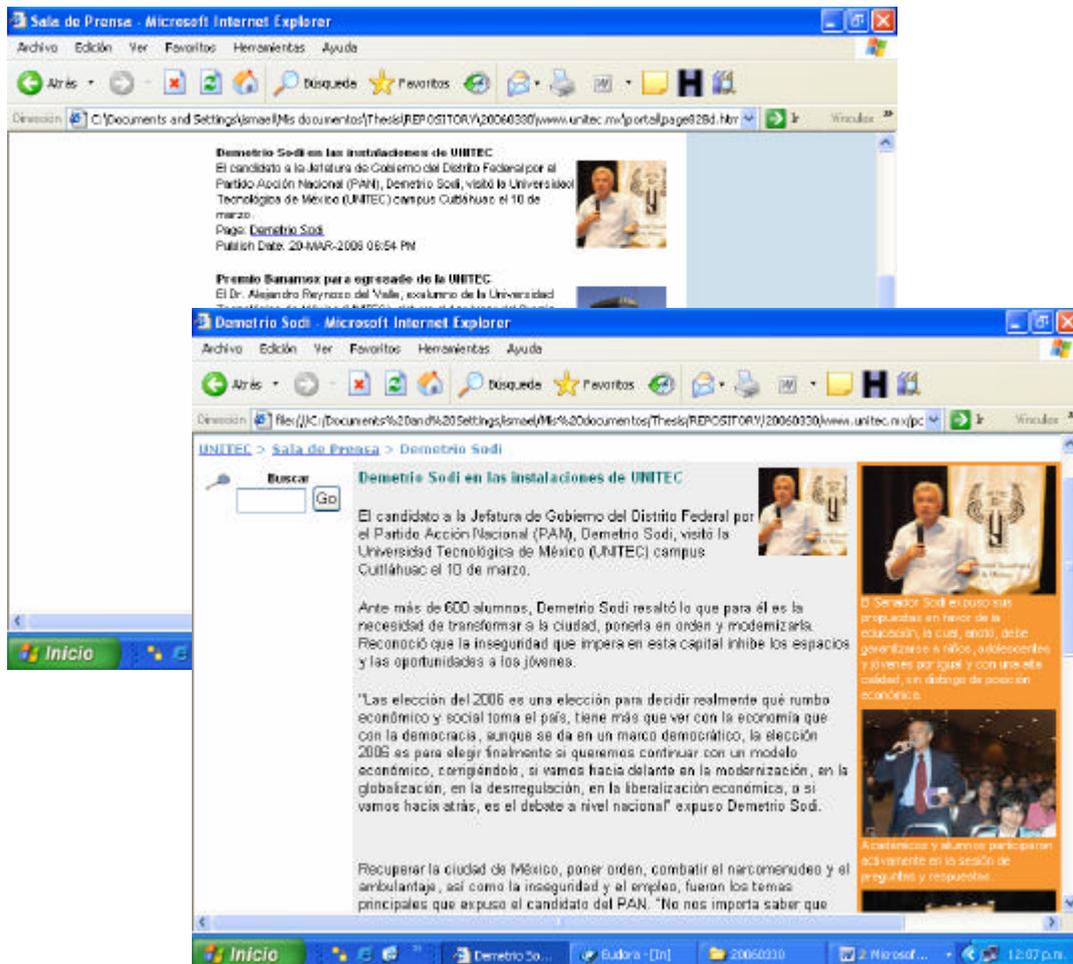


Figura 5.4. Una lista de noticias con encabezado, una descripción parcial y un enlace a otra página complementaria.



Figura 5.5. Una relación de encabezado y descripción completa del evento

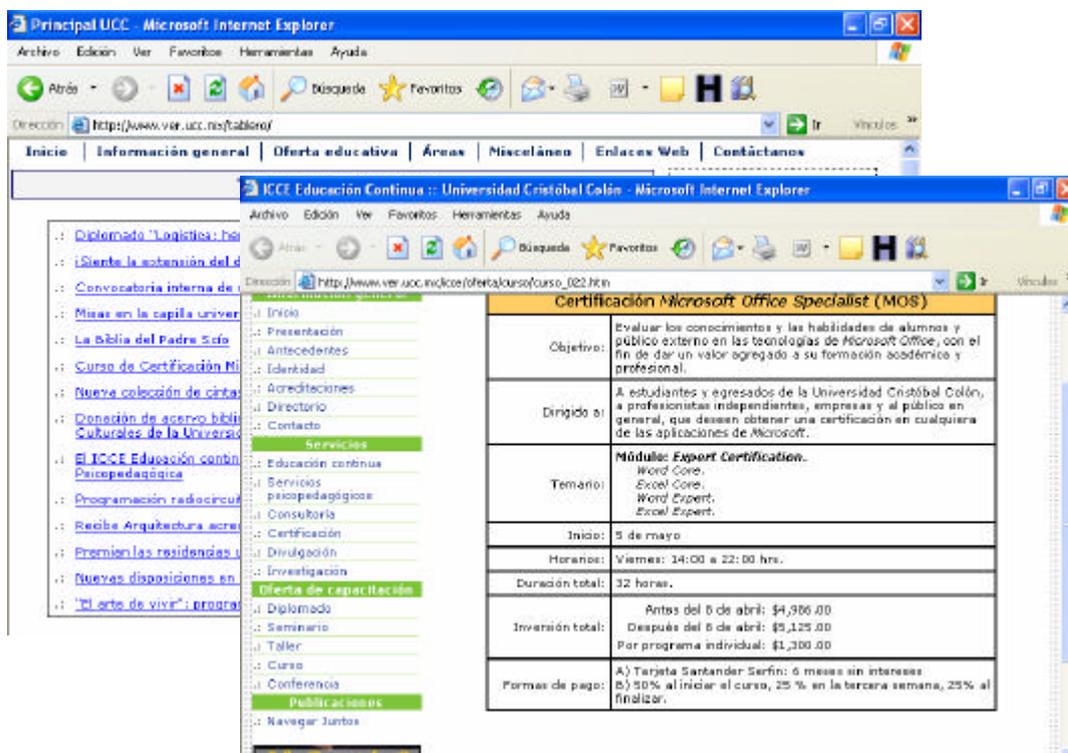


Figura 5.6. Un listado de títulos de eventos con enlace hacia páginas que informan a detalle

5.2 SELECCIÓN DE LAS EMPRESAS A ANALIZAR

El trabajo a desarrollar en el presente estudio, tiene como meta obtener un proceso automatizado aplicable a cualquier sector o área industrial, que sirva a las necesidades de sistemas de inteligencia empresarial y/o competitiva; sin embargo, para fines demostrativos se ha enfocado al sector educativo de nivel superior, por la naturaleza de sus publicaciones, como se describe enseguida.

Las universidades presentan actividades periódicas cíclicas para responder a las necesidades de instrucción y formación de los usuarios. Necesitan plantear con antelación, las actividades de sus programas de estudio y comunicarlas por distintos medios y en los momentos más adecuados.

Tienen entre sus actividades sustantivas, además de la educación, la difusión de la ciencia y la cultura, convirtiéndose en uno de los foros por excelencia, para allegar y comunicar eventos de carácter cultural a usuarios internos y externos.

El desarrollo de habilidades y destrezas en alumnos, maestros y administrativos, requieren orientarse a las necesidades que plantea el mundo empresarial, dando un gran énfasis a las competencias de valoración objetiva y susceptibles de certificarse por organismos externos a la casa de estudios. De ahí que, las universidades tienen la tarea diaria de difundir los programas de capacitación que el mercado requiere.

La formación del alumnado para ser integral, requiere de hacerle vivir su experiencia universitaria más allá de las aulas y laboratorios, necesita fomentar actividades de socialización y deportivas. La universidad en su afán de lograr una mejor imagen pública, reseña estas actividades con lujo de detalles en textos, gráficas y video, aprovechando las ventajas del Internet.

Para elevar el desarrollo regional donde reside la universidad, se realizan actividades de vinculación entre los sectores empresarial y educativo, que tienen mayor resonancia para ambas entidades, cuando se hacen del conocimiento del público, de la manera más adecuada.

Como puede observarse, las necesidades de comunicación por diversos medios, de preferencia Internet, son fuertes, constantes y altamente productivas para el sector universitario.

En el país, este sector se divide por razones económicas, en públicas y privadas. Éstas últimas por lo general, desarrollan múltiples actividades tendientes a allegarse mayores fondos, para sustentar su operación e inversiones. A diferencia de sus homólogas, de carácter público, necesitan informar continuamente sobre aquellas actividades que las hagan más atractivas y lograr convertir a dichos públicos en fuentes de financiación. Los fondos así obtenidos, han de aplicarse a cubrir tanto necesidades de la universidad como de las entidades que la misma beneficia con fines altruistas.

Para las privadas, es muy importante diversificar y atraer la mayor cantidad de fuentes de fondeo y una de las estrategias más usuales, es lograr una imagen favorable a través de la comunicación constante de sus avances y logros. Por estas razones, para fines del presente estudio, se considera solamente a las universidades privadas, de las cuales se ha considerado a veinte de ellas:

ANAHUAC	LA SALLE BAJIO	UIC
CETYS MEXICALI	TECMILENIO	UNITEC
CUGS	ITESM, CAMPUS PUEBLA	UVM
EBC	UAG	UNIV. PANAMERICANA
IBERO CD. MEXICO	UAM	UNIV. REGIOMONTANA
ITESO	UCC	UPAEP
LA SALLE	UDEM	

Tabla 5.1. Las universidades del sector privado.

El criterio de selección se basó en la riqueza en cuanto a cantidad y calidad de sus publicaciones, con la finalidad de generar una solución aplicable a cualquier otro sector industrial. Estas empresas se consagran además en el anexo A.

5.3 SELECCIÓN DEL MÉTODO DE TRANSFORMACIÓN

El desarrollo de los métodos para la transformación de contenido textual ha sido paralelo al desarrollo de los métodos para la comprensión del lenguaje (análisis morfológico, sintáctico y semántico). Por ello, típicamente se busca y analiza la información textual considerando únicamente el “tema” de los textos y no su contenido completo. Esta estrategia facilita el análisis de grandes conjuntos de textos, e incluso mantiene una independencia del dominio, pero limita grandemente la expresividad y la diversidad de los resultados de los sistemas de análisis de textos.

Para elegir el método de transformación aplicado en la presente investigación, se presenta de forma sucinta las principales corrientes actuales.

Montes [51] establece que, tradicionalmente la minería textual analiza las relaciones entre pequeñas unidades que expresan sólo aspectos temáticos del texto (acerca de que ó el quién del texto). La idea es analizar la información que puede obtenerse de estructuras mayores (frases u oraciones representadas por grafos conceptuales), las cuales preservan mucho más de la semántica del texto y permiten construir jerarquías conceptuales a través de la descripción numérica y simbólica de su similitud.

Sin embargo, como lo indica Boytcheva [52], la extracción automática de especificaciones de conocimiento formal desde texto en lenguaje natural es una tarea ardua, pues la adquisición automática de conocimiento necesita de una descripción previa de cada palabra importante, esperada en el texto de entrada. Una palabra puede ser reconocida y luego procesada si y solo si puede ser derivada desde una entrada léxica aplicando reglas morfológicas. Para que sea un proceso automático requiere: definiciones de miles de palabras, esfuerzos para proveer análisis de texto completo y atención especial a la consistencia del conocimiento, el cual es adquirido y almacenado en la base de conocimiento (KB). Debido a todas esas complicaciones, la automatización actualmente, es posible solamente para lenguaje natural restringido.

Así por ejemplo, Rassinoux y otros [53] al dedicarse al desarrollo de un analizador multilingüe, denominado RECIT, orientado a la comprensión de textos

provenientes de documentos de pacientes del departamento de cirugía digestiva, determinan que el estilo idiomático encontrado en estos textos conforma una estructura simple y restringida en comparación con el lenguaje natural completo. Con ello, se puede implementar un sistema que trata automáticamente con expresiones importantes sólo para el dominio médico en esa específica área.

Hess [54] indica que, como el vocabulario en un dominio restringido, como el descrito en su trabajo sobre dispositivos digitales, es más bien limitado, el número de grafos de palabras con sentido que debe agregarse para analizar cada instrucción, decae conforme el diccionario alcanza el límite. Se puede esperar que el número de grafos por frase decaiga hacia el final de cada documento dado que los autores tienden a repetirse en su estilo.

Ahora bien y de acuerdo con Petersen [55], existen dos clases de métodos para la extracción automática de estructuras conceptuales desde textos. Una de ellas está más inclinada hacia un manejo ontológico y menos hacia un manejo sintáctico y está basado en un inventario de grafos denominados canónicos, organizados en un árbol. Ejemplos de trabajos hechos con esta clase de metodología incluyen los de Sowa y Way [56], Sowa [57] y Hess y Cyre [58].

La otra clase de método está menos orientado a la ontología y más hacia la sintaxis. Está centrado alrededor de la conversión de estructuras sintácticas a estructuras conceptuales, usando grafos intermedios, reglas de conversión, apareamiento de grafos, reunión de grafos, selección de relaciones conceptuales basado en la colocación de tipos de conceptos en la ontología y posiblemente en la heurística e información estadística. De este lado, se encuentran trabajos de Barrière [58]. Existe además una variante que combina ambas propuestas para reducir las debilidades inherentes (Boytcheva et al., 2001).

Las noticias empresariales publicadas en la Web es un caso más de textos de dominio específico, razón por la cual se eligió una estrategia basada más en la sintaxis.

La idea central que sustenta el procedimiento propuesto es transformar los textos sin estructura de las noticias a una representación estructurada de archivos de base de datos relacional, susceptible de explotarse mediante consultas programadas o ad-hoc.

De modo que de un texto como:

Tras diez meses de ausencia, los alumnos de Administración de Instituciones de la Universidad Popular Autónoma del Estado de Puebla (UPAEP), Paola Cortés Paz, Diana López Cruz y Rodrigo Tello Rosete, retornaron a esta casa de estudios después de permanecer en Tesalónica, Grecia, como parte del proyecto “Desarrollo Sustentable del Turismo entre la Unión Europea y Latinoamérica”, auspiciado por el programa América Latina-Formación Académica (ALFA).

Se obtenga una estructura, así:

Elemento	Contenido
Lugar	NA
Fecha	NA
Agente	los_alumnos_de_Administración_de_Instituciones_de_la_Universidad_Popular_Autónoma_del_Estado_de_Puebla_(UPAEP), _Paola_Cortés_Paz, Diana_López_Cruz y Rodrigo_Tello_Rosete,
Accion	Retornar
Objeto	
	Rf = Referente Casa
	Drf = Determinativo del ref. de_estudios
	C = Calificación NA
	Dc = Determinativo de la Cal. NA
Paciente	NA
Complemento	
	I = Instrumento el_programa_América_Latina-Formación_Académica_(ALFA).
	E = Efecto NA
	K = Contexto del_proyecto_“Desarrollo_Sustentable_del_Turismo_entre_la_Unión_Europea_y_Latinoamérica”,
	Relator NA

A la cual se pueda leer mediante el lenguaje universal de consultas a bases de datos: SQL.

5.4 SELECCIÓN DEL ESQUEMA DE TRATAMIENTO

El pre-procesamiento en la minería textual, se encarga de transformar los textos sin estructura a contenidos intermedios semi-estructurados o estructurados, con el fin de aplicar diversas técnicas de análisis cuantitativo y obtener el conocimiento meta. De los distintos métodos revisados para el tratamiento de noticias, susceptibles de aplicarse, el trabajo del Dr. Colle [50], consistente de un análisis lógico-semántico de noticias, fue seleccionado para automatizarse, por las enormes posibilidades de su aplicación a la minería textual.

De acuerdo con Colle [50], la noticia implica por definición la existencia de un Hecho y conforme al modelo básico de su estructura, se detectan habitualmente sus componentes, en distinto orden; actor, acción, objeto y paciente:

- el actor o sujeto-agente, que realiza la acción principal
- la acción, que corresponde al verbo principal del núcleo noticioso
- el paciente, quien sufre las consecuencias de la acción o es el destinatario directo de ella
- el objeto de la acción, ya sea tangible ó intangible
 - tangible: la realidad material a la cual se aplica la acción (ej. la bala que se disparó, el edificio que se incendió, la comida que se comparte)
 - intangible: el contenido de la misma (ej. tema de una declaración).

Este esquema básico puede tomar cinco aspectos diferentes (3 basados en tangibles y 2 en intangibles), a partir de los cuales es posible construir formas más complejas:

Basados en tangibles:

1. X hace algo. En este caso no hay afectado, excepto el mismo X, a veces
2. X afecta a Y
5. X entrega algo a Y

Basados en intangibles:

1. X dice algo a Y
2. X dice algo a Y acerca de un hecho anterior

En ocasiones, es posible agregar al hecho esencial, la persona que relata el evento, esto es, el relator, a partir de lo cual, se enriquecen las formas anteriores.

Así mismo, el lugar y fecha, son gramaticalmente "complementos circunstanciales"; desde el punto de vista intrínseco, no son datos complementarios ni tampoco forman parte de la descripción propia del acontecimiento; pero son referencias esenciales del contexto espacio-temporal, indispensables para ordenar y valorar la información. De este modo, la fórmula básica de la reseña se transforma en Lugar, Fecha, Hecho y Relator.

Por su propia naturaleza, el objeto puede componerse de dos partes básicas:

- El referente (Rf), que corresponde a la entidad implicada (otro acontecimiento, un objeto material o un concepto, como el "tema" de una declaración).
- La calificación (C), que corresponde al juicio que formula el actor acerca del referente (o de la acción, si su definición incluye referente).

Sin embargo, para el caso de hechos tangibles, no forzosamente aparecen estos elementos. Ejemplos:

- Los activistas atacaron dos policías (Sin objeto)
- Bomberos apagan incendio de supermercado (Con objeto, pero no calificado)
- El embajador regresó sorpresivamente a su país (Sin objeto con calificación)

Para los hechos intangibles, el objeto ha de manejar el referente cuando menos. Ejemplos:

- La Junta aprobó una ley. Objeto = la ley

- El ministro del Interior calificó de criminales despiadados a quienes colocaron una bomba en el consultorio. Objeto referente = una colocación de bomba (otro acontecimiento), calificación = acto criminal
- El secretario general del PDC calificó de malévolas las declaraciones que atribuyen a su partido un acuerdo secreto con el PC. Referente = declaraciones anteriores; calificación : malévolas.

La calificación (C) algunas veces aparece implícita en el verbo de acción como en "compartir un planteamiento" que conviene traducir por "expresar su acuerdo".

Además pueden aparecer determinativos del objeto (D), los cuales representan todos los términos que complementen la definición del objeto, detallando más claramente su contenido o sentido. Pueden aparecer determinativos tanto para el referente como para la calificación. Ejemplo.

"Punta Arenas, .././1987. El obispo planteó que la tarea del gobierno consiste en proponer al país un camino coherente...":

Objeto (referente) = la tarea; determinativo = (consistente en) proponer un camino coherente; calificación: no explícita / implícitamente: es un deber

5.5 SELECCIÓN DE LOS ELEMENTOS POR PROCESAR

Como se ha referido anteriormente, toda noticia tiene tres elementos distintivos: el título, primer párrafo (*lead*) y párrafos adicionales. Sin embargo, la parte sustantiva del hecho que se desea notificar se encuentra en el título y en el primer párrafo. A partir de estas nociones, en el presente trabajo se procede a procesar solamente los componentes referidos, para lo cual se plantean algunas consideraciones.

Títulos

Los títulos de noticias contienen elementos distintivos que permiten el procesamiento de parte o la totalidad del hecho noticioso. De acuerdo al estilo de la entidad comunicadora, hay diversas expresiones de titulares, de los breves e inexpressivos hasta los totalmente explícitos:

- *Diplomado*
- *Centro de Lenguas*
- *FIDELIDAD INSTITUCIONAL*
- *Hay que tener verdadera voluntad de escribir*
- *EXPOSICIÓN DE PAPEL MACHÉ “ALEBRIGES Y CATRINAS”*
- *FINALIZAN DIPLOMADO EN COCINA MEXICANA A TRAVÉS DE LOS SIGLOS*
- *Nuestra Universidad fue sede del 4º Encuentro Iberoamericano de Protección de Datos Personales*
- *Inauguran en la UDLA el XVI Congreso Internacional de Electrónica, Comunicaciones y Computadoras*
- *Profesores del Centro de Negocios Agropecuarios Internacionales (CENAGI) recibieron la acreditación de Financiera Rural*

Es necesario referir que existen situaciones especiales, mismas que se describen enseguida:

- En ocasiones el primer párrafo complementa el título, como puede apreciarse enseguida.

Diplomado

En Derechos Humanos para periodistas y profesionales...

- En otras, el título corresponde al determinativo del objeto y generalmente se inicia el primer párrafo con términos como: *Sobre este tema, sobre esta temática*, etc.

Ejemplo.

Capacitación en técnicas persuasivas para los diseños

Sobre esta temática versó el curso de capacitación que impartió hace algunos días el Dr. Román Esqueda.

- Ocasionalmente el titular posee todos los elementos que se desean conocer de la noticia, en cuyo caso se ahorra el tratamiento del encabezado.

Encabezados (leads)

El encabezado generalmente se distingue por ser el primer párrafo, aunque cuando se refiere a un hecho intangible, esto es, cuando se trata de una declaración o aseveración, el encabezado requiere del párrafo siguiente además del título, donde regularmente se encuentran algunos elementos básicos.

Para proceder a su procesamiento, se debe cumplir cualquiera de los siguientes casos:

- Cuando no se encuentran alguno de los elementos básicos: verbo, objeto y agente ó paciente.
- Para revisar si existe uno o varios textos delimitados por comillas, lo cual conduce a considerarlo como determinante del objeto y por tanto enriquecer los elementos base.

- Ocasionalmente en el título se menciona el agente por sus siglas para abreviar el título. En el primer párrafo sin embargo, es posible encontrar el área específica del mismo.
 - Ejemplo.

UPAEP a las semifinales de fútbol soccer.

Las Águilas de la Universidad Popular Autónoma del Estado de Puebla –UPAEP- lograron clasificarse a la Fase de Semifinales.

5.6 TRATAMIENTO MANUAL DE UNA MUESTRA DE TITULARES DE NOTICIAS

A partir de la propuesta anterior, se efectuó un análisis manual de las noticias publicadas el día 17 de noviembre de 2005, en los portales de las empresas escogidas para la muestra. Se realizó este ejercicio, para detectar algunos puntos representativos en las noticias publicadas y delinear un patrón que sienta las bases para la construcción del algoritmo computacional.

Este primer procesamiento se aplicó solamente a los titulares para determinar el grado de efectividad de la solución automatizada y se encontró que:

- Los eventos notificados, en algunas ocasiones, corresponden a las acciones (verbos), pero en otras, a los objetos
- Solamente aquellos titulares con un evento claramente definido, fueron tratados
- Por su naturaleza, el lugar y fecha no aparecen en las noticias empresariales, pero la fecha corresponde al día que se visitó al sitio y éste, al lugar
- En su gran mayoría, el actor de manera implícita, es la empresa que publica las notas
- De manera similar, el paciente generalmente no se indica, pero atañe a la comunidad a la que sirve
- Gran parte de los titulares aceptados, contiene el referente y determinativo del objeto, más no así, la calificación y determinativo correspondiente.

Conforme se fueron procesando los titulares, se detectó que varios no indicaban acción ni objeto, razón para no considerarlos y determinar que para esos casos, se requiere un análisis más amplio sobre el encabezado ó primer párrafo de la nota. En raras ocasiones, algunas noticias manejan solamente el título y el mismo no se pudo procesar, indicativo de una incorrecta redacción.

Una vista parcial de los resultados del tratamiento del lead, aparece en la tabla 4.

Actor	Acción	Objeto			Paciente
		Referente	Determinativo	Calif. Det.	
Mundo Ejecutivo	Posiciona	Programa de Maestría	en Alta Dirección		Univ. Anahuac
IDEA	Analiza	evolución y perspectivas			economía mexicana
Anahuac	Recibió	Medalla Anáhuac	en Relaciones Internacionales 2004		Embajador Emérito Sergio González
Alumnos y líderes del turismo	Se unen	combatir	Pobreza		en Mexico
UA	celebra	congreso	"Actualidad en el Deporte Mexicano"		
CETYS Universidad	realizó	premiación	Programa Jóvenes Emprendedores		
Samsun	otorgó	becas			estudiantes de CetyS Universidad
CETYS Universidad	clausura	taller	Pintura		estudiantes de profesional
CETYS Empreses	presentó	4 casos		de éxito	
CETYS Universidad	realizó	final	Simulación de empresas		
CETYS Universidad	reunión			exitosa	exalumnos
Egresados	develan	placa			
Ex alumno	imparte	conferencia			
C.P. Rodolfo Pinedo Torres	exposición	Máquinas de oficina	una herramienta del contador		
EBC	presentación	Lic. En Admon. Hotelera y Turística	Modalidad Nacional		
EBC	entrega	cartas	de Candidatura		
EBC	entrega	becas	Excelencia académica		
EBC	presentación	Lic. En Admon. Hotelera y Turística	Modalidad Nacional		
EBC	celebra	día	del Maestro		
Ibero Cd. De México	expone	proyectos	fotográficos		
Quórum y Canal Once	premian	Excelencia	Académica		Diseño
ACEUJ	organiza	conferencia			Egresados
Biorretroalimentación	ayuda	a prevenir	cuadros		Estrés
Iteso y Ciatej	firmaron	acuerdo	constituir un biocluster		
Iteso	brindó	apoyo	en combate		al incendio
Académicos	reciben	reconocimientos	en los Premios Psicología		

Tabla 5.2. Vista parcial de los resultados

De los verbos o acciones más representativos encontrados en este análisis, fueron:

aceptar	engalanar	planear
analizar	entregar	planificar
anunciar	ser	posicionar
apoyar	estudiar	premiar
ayudar	exponer	presentar
brindar	firmar	realizar
buscar	ganar	recibir
celebrar	graduar	reconocer
clausurar	hablar	reunir
concluir	impartir	rendir
consolidar	inaugurar	titular
convenir	iniciar	unir
crear	llevar	seleccionar
despedir	obtener	tomar
develar	organizar	trabajar
difundir	otorgar	venir

Tabla 5.3 Verbos más representativos

Para el caso de los objetos, se encontró que algunos corresponden a verbos, con una de las siguientes terminaciones: ción, sión y miento; tal fue el caso de:

celebración	planeación
conclusión	premiación
exposición	presentación
inauguración	recepción
impartición	rendición
iniciación	recibimiento
obtención	reunión
organización	reconocimiento

Tabla 5.4 Objetos especiales

A partir de lo anterior, se presenta la relación de objetos más representativos de los diferentes tipos de evento que la universidad notifica.

acceso	coctel	doctorado	juramento	proyecto
actuación	coloquio	encuentro	libro	recepción
acuerdo	competencia	escuela	licenciatura	recibimiento
alianza	concierto	estudios	maestría	reconocimiento
apoyo	conclusión	exhibición	medalla	rendición
asamblea	concurso	exposición	noche	reunión
beca	conferencia	feria	nombramiento	semana
brindis campamento	congreso convenio	festival fiesta	obtención organización	seminario servicio
capacitación	convivencia	final	placa	simposio
carta	convocatoria	firma	planeación	simposium
cátedra	curso	foro	potencia	taller
celebración	desayuno	graduación	premiación	temporada
cena	despedida	impartición	premios	titulación
ceremonia	Día	inauguración	presentación	toma de protesta
certificación	diplomado	iniciación	programa	videoconferencia
				vinculación

Tabla 5.5 Objetos más representativos

Con los datos recabados de esta revisión manual, se puede determinar que, para conocer lo que en términos generales hace la universidad, es aplicable el procedimiento propuesto y a partir de ahora, se está en camino de automatizarlo.

Adicionalmente, los objetos y verbos detectados serán la base de los conjuntos de términos especiales por utilizarse más adelante, los cuales son consagrados en los anexos B y C, respectivamente.

5.7 CORPUS UTILIZADO

Términos conectores y delimitadores del idioma

Los términos del idioma que conectan o delimitan (artículos, adjetivos, preposiciones, conjunciones, adverbios y verbos auxiliares) a los términos sustantivos de una oración (nombres propios, nombres comunes, verbos), servirán para el análisis sintáctico de los títulos y/o encabezados. A partir de los consagrados en el anexo D, se determina más adelante, aquellos términos que sirven de interconexión ó delimitación y la manera en que se usarán.

Los elementos usados en el presente trabajo, provienen del trabajo de Sebastián y otros [59] denominado “LEXESP: Léxico Informatizado del Español”, resultado de un proceso de anotación y etiquetado morfológico de millares de términos provenientes de dos corpus: CliC-TALP y LexEsp. El primero proviene de dos muestras: Una muestra representativa (de 500,000 palabras) de un corpus de prensa de 7 millones de palabras cedido por el periódico La Vanguardia; y otra, el cual consta de 5 millones de palabras, recogido entre los años 1978 y 1995.

Luego de una revisión manual de varias noticias, se ha encontrado la enorme ambigüedad que presentan las preposiciones, artículos y conjunciones. Así por ejemplo, la preposición *de* puede usarse como indicador de posesión, origen, pertenencia, cantidad, herencia, mientras que *por* puede asociarse a duración, razón, medio, trayecto y otros. Ante esto, se requiere determinar cuales términos están asociados a los elementos básicos que se pretenden buscar, para eliminar dicha ambigüedad.

Así por ejemplo, uno de los elementos básicos, es la entidad que realiza la acción principal que se reporta en la noticia, quien recibe el nombre de Agente y es representado por un nombre propio. Es posible detectar los nombres propios porque la primera letra es mayúscula generalmente. Para determinar el agente a partir de un nombre propio, se buscan algunos términos que pueden antecederle, como:

al, como, de acuerdo a, de acuerdo con, nuestra(o), por

Más aun, el agente puede componerse de varios nombres propios o un sustantivo y un nombre propio. Las palabras que los conectan pueden buscarse entre:

*a cargo de, a través de, a través del, como, con, conformado
por, de, del, e, en, nuestra(o), para, perteneciente a,
que dirige, y*

La misma tarea se ha realizado para los demás elementos básicos de una noticia, mismos que se consagran en el mapeo de elementos básicos (anexo E).

Las locuciones no verbales.

La anotación que realizaron Sebastián y otros en su trabajo, no es estrictamente palabra a palabra ya que se considera la existencia de locuciones adverbiales, preposicionales y conjuntivas que se etiquetan de modo equivalente a los adverbios, preposiciones y conjunciones. De acuerdo con Ruiz [60], las locuciones son expresiones formadas por grupos de palabras que equivalen a conjunciones, adverbios, preposiciones, adjetivos y verbos. Especialmente para el caso del adverbio, se encuentra que su respectiva locución (en consecuencia, de piedra, al lado, etc.) está compuesta de un número considerable de elementos que no se ajustan a los criterios establecidos por la gramática. Las locuciones coadyuvan a un procesamiento más sencillo ya que engloban y permiten la abstracción de varios términos continuos. A partir del trabajo referido en el primer párrafo, se organizan y presentan en el anexo F.

5.8 PROCEDIMIENTO PARA EL TRATAMIENTO DE LAS NOTICIAS

Por cada noticia, se llevan a cabo dos acciones complementarias, se procesa el título y se trata el primer párrafo.

A partir del trabajo de Esquivel [61], se ejecutan tres etapas: Configuración, preprocesamiento y post-procesamiento.

Configuración.

Inicialmente se almacenan los objetos y verbos descubiertos del procesamiento manual. Periódicamente se agregan verbos y objetos nuevos al registro que se describe en el anexo B y C, provenientes de los elementos detectados y guardados durante la fase de lectura que más adelante se describe.

Además, de vez en cuando se registran y/o modifican datos de la empresa competidora de interés, su nombre y dirección Web donde se publican sus avisos. La relación de empresas se consagra en el anexo A.

Las actividades en esta etapa modifican el contenido de las tablas EMPRESAS, REFERENTES Y ACCIONES, mostradas en el modelo de conjuntos de entidades y su relación de la figura 5.7, para una base de datos relacional.

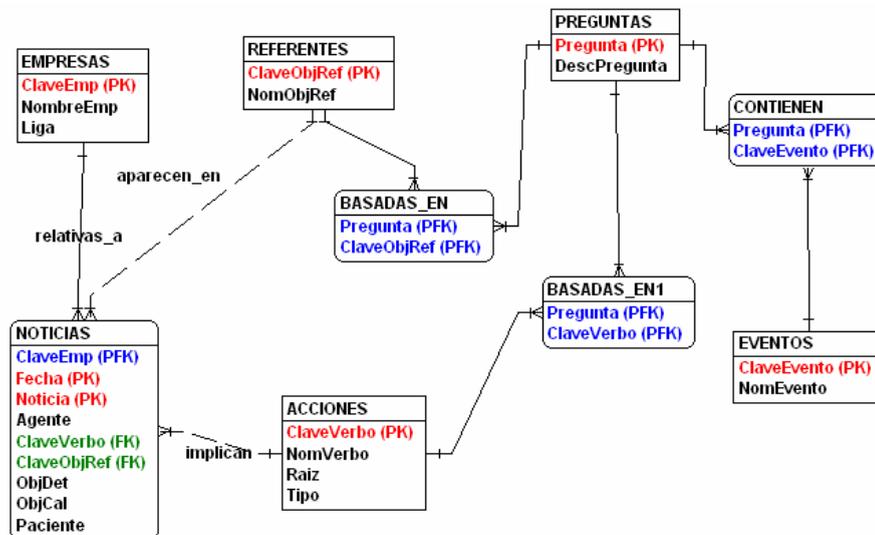


Figura 5.7 Diagrama ERM

Dicha base de datos, contiene los elementos de referencia y por registrar, una vez que se realizan las tres etapas antes señaladas. Para una mejor comprensión, se describen en la tabla 5.6 los archivos que conforman la base de datos y en el anexo I, se documenta el diccionario de los datos.

Archivo	Descripción
NOTICIAS	Contiene la noticia en formato estructurado
EMPRESAS	Maneja las empresas por procesar
REFERENTES	Mantiene los referentes de objetos contenidos en la noticia
ACCIONES	Comprende los verbos principales a detectarse en la noticia
PREGUNTAS	Engloba las preguntas a ser respondidas para fines de la inteligencia competitiva
BASADAS_EN	Registra los referentes asociados a preguntas genéricas que permitirán las consultas al sistema, acciones y
BASADAS_EN1	Registra las acciones asociadas a preguntas genéricas que permitirán las consultas al sistema
EVENTOS	Enlista los hechos que comprende cada una de las preguntas genéricas
CONTIENEN	Permite la posibilidad de que varios eventos se asignen a más de una pregunta

Tabla 5.6 Descripción de los archivos de la base de datos

Preprocesamiento.

Se efectúan operaciones de lectura (figura 5.8) y escritura (figura 5.9), en una base periódica, como se describen a continuación:

Lectura

- Se obtiene del registro de empresas, la dirección electrónica específica donde se publican las noticias
- Se obtiene una copia del archivo HTML que contiene las noticias
- Se convierte a archivo texto sin etiquetas
- Se procede a transformar el archivo resultante, eliminando de cada noticia, el texto fuera del título y primer párrafo.
- Se compara el archivo obtenido con una versión anterior, si son iguales es indicativo de que todavía no se han publicado noticias nuevas en el sitio y se procede con la siguiente empresa
- De ser una versión distinta, el archivo que resulta se duplica como versión anterior para su posterior comparación. El nombre del archivo se agrega al listado de empresas por trabajar
- Se comparan los títulos obtenidos con los de la versión anterior y en caso de coincidir alguno de ellos, se elimina del archivo de trabajo.

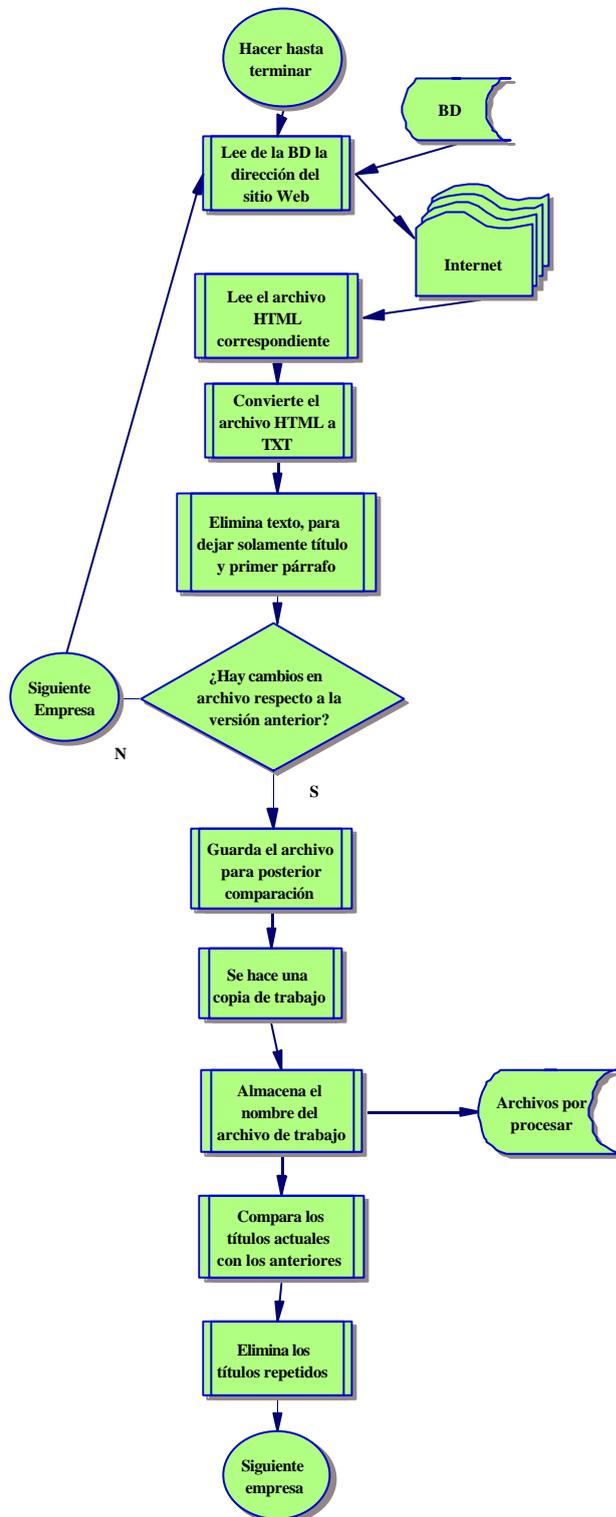


Figura 5.8 Pasos del preprocesamiento (Lectura)

Escritura

Del listado anterior, se realizan las siguientes operaciones:

- Se obtiene el nombre del archivo por procesar
- Se lee del archivo correspondiente, cada uno de los títulos que la conforman
- Por cada título
 - o Se verifica que las palabras del título contengan solo mayúsculas ya sea en la primer letra o en todas
 - o En caso afirmativo, se procede a convertir todas a minúsculas
 - Se descartan términos no significativos (artículos, preposiciones, conjunciones, adverbios)
 - Luego en los términos resultantes, se buscan palabras correspondientes a objetos del archivo correspondiente (Anexo B)
 - Al coincidir, se almacena para el posterior tratamiento del encabezado
 - En caso contrario, se verifica que exista en la lista de términos no referenciados. De ser así se incrementa su frecuencia y si no se registra en la lista. Esto permitirá posteriormente una revisión manual para enriquecer la lista de objetos
 - o En caso contrario, el título se procesa con el algoritmo que más adelante se presenta (Máxima Abstracción)
 - Se verifica la detección de los elementos principales (Agente, Verbo y Objeto)
 - En caso positivo, termina el procesamiento de la noticia (postprocesamiento) y se continua con el siguiente titular
 - Caso contrario, se procede al procesamiento del primer párrafo (lead) con el algoritmo denominado Máxima Abstracción

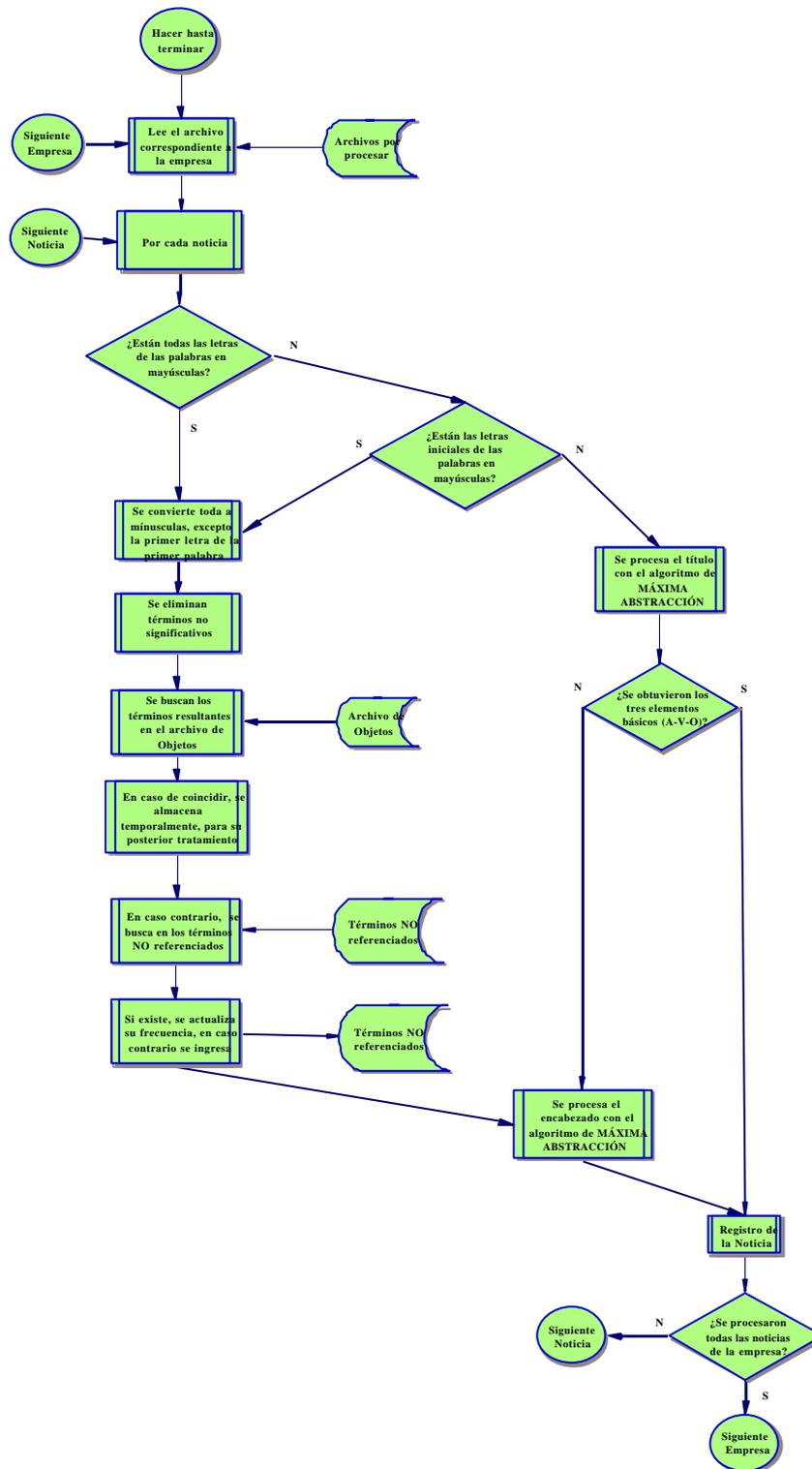


Figura 5.9 Pasos del preprocesamiento (Escritura)

- Se procede a registrar de manera estructurada la noticia (postprocesamiento)
- Se determina si ya se procesaron todos los títulos de la empresa
 - De no ser así, se continua con el siguiente
 - Caso contrario se procede a la siguiente empresa

Post-procesamiento

Por cada noticia procesada, se realizan tareas de ajuste como a continuación se presenta en la figura 5.10 y se describe.

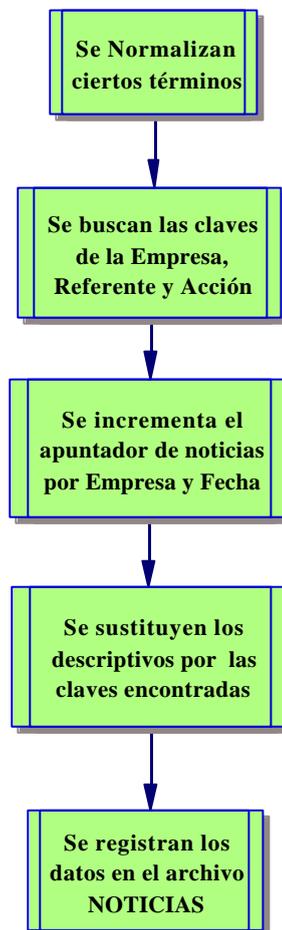


Figura 5.10 Etapa de Post-procesamiento

- Se normalizan los elementos repetibles como son: referentes de objeto, acciones y calificadores de objeto.
- Una vez hecho lo anterior, se procede a buscar las claves numéricas que los representan dentro de algunos de los archivos previamente presentados (REFERENTES, ACCIONES y EMPRESAS).
- En caso de no encontrar ciertos valores, sobretodo para referentes y acciones, se les asigna un valor de cero, que serán referenciados como valores no disponibles (N/D) en su primera ocurrencia.
- Se incrementa el contador de noticias por cada emisión de la empresa para una fecha dada.
- Se procede al registro de manera estructurada de los elementos básicos de la noticia.

5.9 ALGORITMO PROPUESTO (MÁXIMA ABSTRACCIÓN)

5.9.1 CONSIDERACIONES PREVIAS

El objetivo es determinar a partir del primer y segundo párrafos, los elementos básicos del hecho noticioso. Como se ha descrito anteriormente, estos corresponden a:

- el agente (quien realiza la acción principal)
- la acción (que corresponde al verbo principal del núcleo noticioso)
- el paciente (quien sufre las consecuencias de la acción)
- el objeto de la acción
 - tangible (la realidad material que sufre la acción)
 - intangible (el contenido de una declaración)
 - y compuesto de:
 - El referente (la entidad implicada: objeto o concepto)
 - La calificación (el juicio que se formula del referente)
 - Determinativo del objeto (los términos que ofrecen más detalle)

Además, pueden aparecer:

- Lugar y fecha (Referencia espacio-temporal del evento)
- Relator (quien describe el hecho)
- Complemento:
 - Instrumento (que comprende el medio usado)
 - Efecto (que indica los resultados a futuro)
 - Contexto (que expresa las condiciones del evento)

Por lo anterior, se hace necesario establecer antes ciertas consideraciones:

OBJETO.

Para facilitar el procesamiento, se requiere que el objeto conserve la palabra que le antecede, si este es un artículo

>> *la tarea*

>> *el oficio*

El referente del objeto puede determinarse en situaciones como:

- Si después del verbo viene asociado al pronombre posesivo: *su* ó *sus*, los términos subsecuentes corresponden al objeto, el cual no necesariamente se encuentra en la relación de objetos.

El determinativo del objeto se puede detectar en alguno de los siguientes casos:

- Cualquier expresión acotada entre comillas que se exprese un tema o contenido.
- En ciertos textos se puede detectar el determinativo del objeto con los términos: fue dado a conocer, fueron dados a conocer, es uno de los objetivos, fue uno de los objetivos, fue el tema, fueron los temas, fueron los & temas, fueron los & & temas. Donde & corresponden a adjetivos calificativos o adverbios.

AGENTE/PACIENTE.

En la voz pasiva el agente que antecede al verbo es generalmente el PACIENTE. Los términos después de la palabra POR, se refieren al AGENTE

>> *Frei fue elegido presidente por los manifestantes*

>> *Los manifestantes eligieron a Frei como presidente*

>> *Un análisis fue presentado por Angélica*

En algunos casos también, lo que antecede al PACIENTE son términos como: *recibió de, obtuvo de, consiguió de, por el, contarán con, fue reconocido con*

>> *La UNITEC recibió de la SEP*

>> *La UNITEC recibió la acreditación por el CONACYT*

Cuando no se tiene indicativo de quién lo recibe (DE), entonces el que antecede es el AGENTE:

>> *GVA y asociados recibió el premio*

Cuando se usa la tercera persona (se entregó, se invistió, se premió), se infiere que el AGENTE es la universidad/empresa:

>> Se invita a las jornadas espirituales

5.9.2 ESTRUCTURAS DEL ENCABEZAMIENTO.

Luego del análisis manual a un conjunto de noticias muestra, se encontró que los elementos del hecho noticioso que aparecen con mayor frecuencia son: Agente (A), verbo (V) y objeto (O). El orden en que aparecen, origina seis posibles combinaciones de estructura que se ejemplifican en la tabla 5.7.

Estructura	Ejemplo
A-V-O	Kevin reparó la PC
O-V-A	La PC fue reparada por Kevin
V-A-O	Reparó Kevin la PC
V-O-A	Reparó la PC Kevin
O-A-V	La PC Kevin reparó
A-O-V	Kevin la PC reparó

Tabla 5.7 Combinaciones de los elementos de mayor frecuencia

Estas estructuras presentan ciertas características que facilitan la identificación de cada uno de sus elementos. Enseguida se describe cada una de ellas.

Agente – Verbo - Objeto.

Es la combinación más usada por los redactores de noticias y en la que el verbo aparece generalmente en tiempo pretérito.

Ej. *El Café Scientique Guadalajara invitó a su segunda charla del año.*

Objeto – Verbo – Agente.

Corresponde a la voz pasiva, en la cual los verbos auxiliares: haber, ser, poder, entre otros, anteceden al verbo principal en modo participio. El agente es detectado porque generalmente le preceden los términos: con, por, de acuerdo con, etc.

Ej. *El libro “Estructura y Dinámica de la Empresa en México”, fue presentado por María de los Ángeles Pozas.*

Verbo – Agente – Objeto.

Se utiliza generalmente al informar hechos intangibles, en los cuales el objeto es el tema o comentario que un agente hizo. Por ello, el verbo puede ser: afirmar, expresar, decir, opinar, acordar, convenir, acordar y otros, el cual va después de un texto delimitado por comillas y/o una coma.

Ej. “Acomodarse o Morir”, dijo Gabriel Morales, en su plática “El claroscuro de la formación en competencias profesionales”

Verbo –Objeto – Agente.

Es parecida en su aplicación a la anterior estructura, con términos distintos para acomodar primero el objeto y enseguida al agente. Entre los términos especiales, se encuentra que el verbo ser o estar aparece solo o acompañando a un verbo en participio.

Ej. En Monterrey ya no se necesitan profesionistas, esa fue la observación de Gustavo Raymundo Gallo.

Por último, las estructuras O-A-V y A-O-V son realmente casos raros en la forma de expresarse hoy día, por lo cual no serán tratadas en el presente trabajo.

5.9.3 DESCRIPCIÓN DEL ALGORITMO

El algoritmo propuesto recibe el nombre de Máxima Abstracción porque está basado en la actividad humana de lectura, que para el caso de las noticias, se realiza de manera rápida, obviando aquellos elementos del texto que resultan inútiles, para los fines de solamente enterarse sobre lo que se informa. Generalmente, se lee el título de la noticia y si resulta de interés alguna palabra, se procede a leer el primer párrafo del texto, del cual solamente buscamos los elementos que permiten responder a los planteamientos de:

- ¿Qué se hizo?
- ¿Quién lo hizo?
- ¿Qué se usó?
- ¿Quién fue afectado?
- ¿De que trató lo que se hizo o dijo?

Esta búsqueda se facilita por la habilidad que posee la mente humana de abstraer ciertos términos, ocultando o enmascarando otros que solamente sirven de soporte. Si se desea leer rápidamente un primer párrafo como el que sigue:

Tras diez meses de ausencia, los alumnos de Administración de Instituciones de la Universidad Popular Autónoma del Estado de Puebla (UPAEP), Paola Cortés Paz, Diana López Cruz y Rodrigo Tello Rosete, retornaron a esta casa de estudios después de permanecer en Tesalónica, Grecia, como parte del proyecto “Desarrollo Sustentable del Turismo entre la Unión Europea y Latinoamérica”, auspiciado por el programa América Latina-Formación Académica (ALFA).

Los elementos correspondientes a artículos, preposiciones, conjunciones, artículos, pronombres, locuciones, participios de presente, verbos de apoyo son obviados, mientras que los nombres propios y comunes así como las acciones principales atrapan la atención del lector. Más aún, esta tarea se facilita, si el redactor escribe en mayúscula la letra inicial o todas de los nombres propios o bien acota entre comillas y/o paréntesis, los términos de importancia.

Basado en este principio, el algoritmo que se propone está compuesto de los siguientes pasos, mismos que se describirán ampliamente y que se soportarán por cuatro ejemplos subsecuentes.

PASOS

1. Encontrar y englobar los nombres propios
2. Etiquetar y sustituir nombres propios
3. Determinar el objeto, en todos y cada uno de sus elementos
4. Detectar las locuciones no verbales y sustituir
5. Encontrar los verbos y sustituir
6. Determinar el verbo principal y los secundarios
7. Distinguir los nombres propios
8. Definir el lugar y fecha del evento
9. Analizar los términos no participantes

Texto original:

Tras diez meses de ausencia, los alumnos de Administración de Instituciones de la Universidad Popular Autónoma del Estado de Puebla (UPAEP), Paola Cortés Paz, Diana López Cruz y Rodrigo Tello Rosete, retornaron a esta casa de estudios después de permanecer en Tesalónica, Grecia, como parte del proyecto “Desarrollo Sustentable del Turismo entre la Unión Europea y Latinoamérica”, auspiciado por el programa América Latina-Formación Académica (ALFA).

1. Encontrar y englobar los nombres propios

Los nombres propios se reconocen por tratarse de palabras cuya letra inicial es una mayúscula. En segundo término, se reconocen como siglas, siglas entre paréntesis ó títulos de empresas como en UCC, UPAEP, etc. Inicialmente se asocian aquellos nombres propios directamente conectados. Para ello, se usa el signo “_”.

Texto resultante:

Tras diez meses de ausencia, los alumnos de Administración de Instituciones de la Universidad_Popular_Autónoma del Estado de Puebla_(UPAEP), Paola_Cortés_Paz, Diana_López_Cruz y Rodrigo_Tello_Rosete, retornaron a esta casa de estudios después de permanecer en Tesalónica,_Grecia, como parte del proyecto “Desarrollo_Sustentable del Turismo entre la Unión_Europea y Latinoamérica”, auspiciado por el programa América_Latina-Formación_Académica_(ALFA).

Nombres propios:

- Tras
- Administración
- Instituciones
- Universidad_Popular_Autónoma
- Estado
- Puebla_(UPAEP)
- Paola_Cortés_Paz
- Diana_López_Cruz
- Rodrigo_Tello_Rosete
- Tesalónica,_Grecia
- “Desarrollo_Sustentable
- Turismo
- Unión_Europea
- América_Latina-Formación_Académica_(ALFA)

Luego se extienden hacia delante y hacia atrás de acuerdo a los conectores de nombres propios que se presentan en el anexo E, conformando términos más aglutinantes.

Texto resultante:

Tras diez meses de ausencia, los alumnos de Administración de Instituciones de la Universidad Popular Autónoma del Estado de Puebla (UPAEP), Paola Cortés Paz, Diana López Cruz y Rodrigo Tello Rosete, retornaron a esta casa de estudios después de permanecer en Tesalónica, Grecia, como parte del proyecto “Desarrollo Sustentable del Turismo entre la Unión Europea y Latinoamérica”, auspiciado por el programa América Latina-Formación Académica (ALFA).

Nombres propios extendidos:

- de Administración de Instituciones de la Universidad Popular Autónoma del Estado de Puebla (UPAEP), Paola Cortés Paz, Diana López Cruz y Rodrigo Tello Rosete,
- Tesalónica, Grecia
- “Desarrollo Sustentable del Turismo entre la Unión Europea y Latinoamérica”
- América Latina-Formación Académica (ALFA).

Se requiere extender al máximo los nombres propios, buscando términos asociados como los artículos determinados y los nombres comunes detectados entre estos artículos y los nombres propios. Por tanto, los términos resultantes son:

- *Los_alumnos_de_Administración_de_Instituciones_de_la_Universidad_Popular_Autónoma_del_Estado_de_Puebla_(UPAEP), Paola_Cortés_Paz, Diana_López_Cruz y Rodrigo_Tello_Rosete,*
- *del_proyecto_“Desarrollo_Sustentable_del_Turismo_entre_la_Unión_Europea_y_Latinoamérica”*
- *el_programa_América_Latina-Formación_Académica_(ALFA).*

2. Etiquetar y sustituir nombres propios.

Los nombres propios detectados se etiquetan y sustituyen en el texto original para facilitar su procesamiento. De lo anterior resulta:

Tras diez meses de ausencia, NP1, retornaron a esta casa de estudios después de permanecer en NP2 como parte NP3, auspiciado por NP4.

NP1 = *Los_alumnos_de_Administración_de_Instituciones_de_la_Universidad_Popular_Autónoma_del_Estado_de_Puebla_(UPAEP), Paola_Cortés_Paz, Diana_López_Cruz y Rodrigo_Tello_Rosete,*

NP2 = *Tesalónica,_Grecia*

NP3 = *del_proyecto_“Desarrollo_Sustentable_del_Turismo_entre_la_Unión_Europea_y_Latinoamérica”*

NP4 = *el_programa_América_Latina-Formación_Académica_(ALFA).*

3. Determinar el objeto, en todos y cada uno de sus elementos

Como se ha precisado, el objeto está conformado por el referente, el determinativo, una calificación y el determinativo de la misma. Para determinar el primero, se trabaja con los términos asociados a objetos encontrados en el apartado

previo de procesamiento de los titulares y consagrados en el anexo B. Una vez localizado el referente, se procede mediante el mapeo de elementos básicos (Anexo E) a detectar los demás componentes del objeto y sustituir en el texto por sendas etiquetas.

Objeto referente [OR] = casa
Objeto determinativo [OD] = de_estudios
Calificación del objeto = No aplica (NA)

Tras diez meses de ausencia, NP1, retornaron a esta [OR,OD] después de permanecer en NP2 como parte NP3, auspiciado por NP4

4. Detectar las locuciones y sustituir

Las locuciones se buscan en el texto, a partir de las referidas en el anexo F y se sustituyen por etiquetas. Esta acción facilita la posterior detección de los verbos: principal y auxiliares.

Tras diez meses de ausencia, NP1, retornaron a esta OBJETO[OR,OD] L1 permanecer en NP2 como parte NP3, auspiciado por NP4

L1 = después_de

5. Encontrar los verbos y sustituir

Para llevar a cabo esta actividad, primero se busca alguna expresión de más de una palabra que implique un verbo, como las señaladas en el anexo G. En caso de encontrarse, se realiza la sustitución en el texto. Luego, se detectan los términos que corresponden a adjetivos y sustantivos con la terminación del participio de presente -ante, -ente, -iente y se enmascaran con el término (&), ya que son inútiles para la detección de verbos. Los citados términos se obtienen de la lista presentada en el anexo H. Enseguida, se buscan en el texto, los verbos del anexo C, a partir de su raíz, y al encontrarse se etiquetan y se sustituyen en el texto.

Tras diez meses de ausencia, NP1, retornaron a esta OBJETO[OR,OD] L1 permanecer en NP2 como parte NP3, auspiciado por NP4

Tras diez meses de ausencia, NP1, V1 a esta OBJETO[OR,OD] L1 V2 en NP2 como parte NP3, V3 por NP4

Verbos encontrados:

V1 = retornaron	(retornar)
V2 = permanecer	(permanecer)
V3 = auspiciado	(auspiciar)

6. Determinar el verbo principal y los secundarios

Para elegir al principal, de entre los varios verbos participantes, cada uno de ellos pasa por un proceso de votación de acuerdo a las reglas consagradas en el anexo L; resultando ganador el verbo con mayoría de votos. Para el presente caso, el verbo con más votos es: retornaron

Tras diez meses de ausencia, NP1, VERBO a esta OBJETO[OR,OD] L1 V2 en NP2 como parte NP3, V3 por NP4

Verbo principal: V1

7. Distinguir los nombres propios

De los nombres obtenidos, es necesario distinguir si se trata de un agente, un paciente, el lugar, un relator, un instrumento, el efecto o el contexto, mediante el mapeo de los elementos básicos (Anexo E). En este caso, se analizan:

El agente en ocasiones inicia el párrafo con un artículo determinado o indeterminado. En otras, es el nombre propio que está situado entre la primer coma del párrafo y el verbo principal. Ambas condiciones operan para el presente caso.

Agente=

los_alumnos_de_Administración_de_Instituciones_de_la_Universidad_Popular_Autónoma_del_Estado_de_Puebla_(UPAEP),
_Paola_Cortés_Paz, Diana_López_Cruz y Rodrigo_Tello_Rosete,

El paciente es el nombre propio que representa la entidad afectada por la acción y generalmente está precedido por una preposición o una preposición y un artículo.

Paciente: No aplica (NA)

El contexto se representa por las condiciones en las cuales se da el evento. Es posible identificarle por los términos previos señalados en el anexo E.

Contexto: del proyecto “Desarrollo Sustentable del Turismo entre la Unión Europea y Latinoamérica”,

El instrumento es la serie de términos por los cuales se identifica al medio usado para lograr el efecto planteado. Para identificarle se buscan términos como los señalados en el anexo E.

Instrumento: el programa América Latina-Formación Académica (ALFA).

El efecto está representado por los términos indicativos de beneficio para el paciente, el cual al ser implícito, se infiere que se trata de la propia empresa. Los términos asociados generalmente al efecto se describen en el anexo E.

Efecto: No aplica (NA)

El relator es el nombre propio asociado a verbos que indican una declaración. Para identificarle, se puede apoyar en los términos que le preceden, tal y como se muestra en el anexo E.

Relator: No aplica (NA)

Tras diez meses de ausencia, AGENTE, VERBO a esta OBJETO[OR,OD] L1 V2 en NP2 como parte CONTEXTO, V3 por INSTRUMENTO

8. Definir el lugar y fecha del evento

Debido a que normalmente no aparece de manera explícita el lugar del evento y la fecha, es posible identificar a la empresa y la fecha de publicación como tales datos.

Lugar y Fecha: No aplica (NA)

Al final del procesamiento, las etiquetas se sustituyen por su contenido y se obtienen los elementos básicos del hecho noticioso planteado inicialmente, como se muestra en la tabla 5.8.

Elemento	Subelemento	Contenido
Lugar		NA
Fecha		NA
Agente		los_alumnos_de_Administración_de_Instituciones_de_la_Universidad_Popular_Autónoma_del_Estado_de_Puebla_(UPAEP),_Paola_Cortés_Paz, Diana_López_Cruz y Rodrigo_Tello_Rosete,
Accion		Retornar
Objeto		
	Rf = Referente	Casa
	Drf = Determinativo del ref.	de_estudios
	C = Calificación	NA
	Dc = Determinativo de la Cal.	NA
Paciente		NA
Complemento		
	I = Instrumento	el_programa_América_Latina-Formación_Académica_(ALFA).
	E = Efecto	NA
	K = Contexto	del_proyecto_“Desarrollo_Sustentable_del_Turismo_entre_la_Unión_Europea_y_Latinoamérica”,
Relator		NA

Tabla 5.8 Elementos básicos del hecho noticioso

5.9.4 EJEMPLO 1. NOTICIA DE TIPO AVO

Procedimiento.

1. Encontrar y englobar los nombres propios
2. Etiquetar y sustituir nombres propios
3. Determinar el objeto, en todos y cada uno de sus elementos
4. Detectar las locuciones no verbales y sustituir
5. Encontrar los verbos y sustituir
6. Determinar el verbo principal y los secundarios
7. Distinguir los nombres propios
8. Definir el lugar y fecha del evento
9. Analizar los términos no participantes

Texto original

El Centro Nacional de Evaluación para la Educación Superior A.C. CENEVAL, prepara el nuevo Examen General de Conocimientos para los egresados de la Licenciatura en Derecho.

1. Se inicia englobando los términos indicativos de nombre propio.

El_Centro_Nacional_de_Evaluación_para_la_Educación_Superior_A.C._CENEVAL, prepara el nuevo Examen General de Conocimientos para los egresados de la Licenciatura en Derecho.

2. Luego de la sustitución y etiquetado de nombres propios, resulta:

NP1, prepara NP2 para NP3

NP1 =

El_Centro_Nacional_de_Evaluación_para_la_Educación_Superior_A.C._CENEVAL

NP2 = el_nuevo_Examen_General_de_Conocimientos

NP3 = los_egresados_de_la_Licenciatura_en_Derecho.

3. Luego se busca inicialmente entre los nombres propios y después en los restantes, un término indicativo de objeto o el texto posterior al pronombre posesivo (su, sus). Para este caso se localiza en NP2 el término examen. Procesando los elementos del objeto se encuentra que:

OBJETO REFERENTE (OR)

Previo	[NC]	extensor	[NC]
<i>El</i>	<i>nuevo_Examen_General_de_Conocimientos</i>		

NP1, *prepara* OBJETO[OR] *para* NP3

OR = *el_nuevo_Examen_General_de_Conocimientos*

Se buscan términos previos al término que corresponde al objeto para determinar algún posible calificativo, que para este caso resulta en

NP1, *prepara* OBJETO[OR,OC] *para* NP3

OC = *nuevo*

4. Del texto se buscan posibles locuciones para su etiquetado y sustitución, las cuales no existen para el presente caso.

5 y 6. Se busca entre los verbos consagrados en el anexo C, aquellos existentes en el texto. Para el presente caso se detecta solamente uno, por lo cual el proceso de votación se ignora y se convierte en el principal.

NP1, V1 OBJETO[OR,OC] *para* NP3

V1 = *preparar*

NP1, VERBO OBJETO[OR,OC] *para* NP3

VERBO = V1

7. Para distinguir los nombres propios, se procesan NP1 y NP3. Para ello se considera el orden de los elementos básicos encontrados hasta ahora. En este caso el orden es V-O, por lo cual el nombre propio NP1 puede corresponder a un agente o un paciente. Enseguida se buscan las locuciones y de no existir, las preposiciones, conjunciones o adverbios que pueden ir antes de un elemento básico.

La preposición *para* se busca entre los términos que anteceden al agente, paciente, efecto, contexto e instrumento. Para este caso, se usa como elemento previo a paciente y efecto. Para diferenciarlos, el efecto generalmente contiene un verbo y no contiene nombres propios. Si se considera a NP1 como el agente, NP3 equivale al paciente, resultando en:

AGENTE, VERBO OBJETO[OR,OC] *para* PACIENTE

AGENTE =

*El_Centro_Nacional_de_Evaluación_para_la_Educación_Superior_A.C._C
ENEVAL*

PACIENTE = *los_egresados_de_la_Licenciatura_en_Derecho.*

8 y 9. No proceden, por tanto el resultado se muestra en la tabla 5.9.

Elemento	Subelemento	Contenido
Lugar		NA
Fecha		NA
Agente		El_Centro_Nacional_de_Evaluación_para_la_Educación_Superior_A.C._CENEVAL
Acción		preparar
Objeto		
	OR= Referente	Examen
	OD = Determinativo del ref.	General_de_Conocimientos
	OC = Calificación	nuevo
	ODC= Determinativo de la Cal.	NA
Paciente		los_egresados_de_la_Licenciatura_en_Derecho.
Complemento		
	I = Instrumento	NA
	E = Efecto	NA
	K = Contexto	NA
Relator		NA

Tabla 5.9 Elementos de una noticia de estructura A-V-O

5.9.5 EJEMPLO 2. NOTICIA DE TIPO OVA

Procedimiento.

1. Encontrar y englobar los nombres propios
2. Etiquetar y sustituir nombres propios
3. Determinar el objeto, en todos y cada uno de sus elementos
4. Detectar las locuciones no verbales y sustituir
5. Encontrar los verbos y sustituir
6. Determinar el verbo principal y los secundarios
7. Distinguir los nombres propios
8. Definir el lugar y fecha del evento
9. Analizar los términos no participantes

Texto original

La formación de un emprendedor eficiente comienza en la familia, los estudios en segundo lugar y las empresas incubadoras después, señaló Roberto Adame Doria, director de ADC Consultores de Negocios el 17 de octubre en el Teatro del Centro de la Comunidad de la UDEM

1.

La formación de un emprendedor eficiente comienza en la familia, los estudios en segundo lugar y las empresas incubadoras después, señaló Roberto_Adame_Doria, director_de_ADC_Consultores_de_Negocios_el 17 de octubre en el_Teatro_del_Centro_de_la_Comunidad_de_la_UDEM

2.

La formación de un emprendedor eficiente comienza en la familia, los estudios en segundo lugar y las empresas incubadoras después, señaló, NP1 director NP2 el 17 de octubre en NP3

NP1 = Roberto_Adame_Doria

NP2 = de_ADC_Consultores_de_Negocios

NP3 = el_Teatro_del_Centro_de_la_Comunidad_de_la_UDEM

3. Se busca un término que represente un objeto, entre los nombres propios y de no encontrarse entre los demás términos, de lo que se desprende que:

Del texto se obtienen los términos que pueden representar el objeto, entre los cuales se encuentran *formación*, los cuales son verificados contra los previos y extensores de objetos (Anexo E), para encontrar que:

OBJETO

OBJETO REFERENTE (OR)

Previo [NC] extensor [NC] [delimitador]
La formación de un emprendedor... ,

OBJETO[OR], *señaló*, NP1 *director* NP2 *el 17 de octubre en* NP3

4. Se buscan los verbos, a partir de su raíz, en el texto

OBJETO[OR], V1, NP1 *director* NP2 *el 17 de octubre en* NP3

V1 = señalar

Al encontrarse sólo un verbo, este es el principal

VERBO = V1

OBJETO[OR], VERBO, NP1 *director* NP2 *el 17 de octubre en* NP3

7.

En este caso, se encuentra que la estructura de la noticia puede ser del tipo O-V-A ó O-V-P, por lo cual, el nombre propio resultante puede ser agente o paciente. Para distinguirlos se aplican los previos y extensores contenidos en el anexo E y de acuerdo con ello, NP1 corresponde al agente.

OBJETO[OR], VERBO, AGENTE *director* NP2 *el 17 de octubre en* NP3

AGENTE = NP1

Es necesario denotar que cuando hay dos nombres propios separados por un nombre común, se propone que el primer nombre propio absorbe a dicha palabra y al segundo. Por tanto, resulta:

OBJETO[OR], VERBO, AGENTE *el 17 de octubre en* NP3

AGENTE = NP1 *director* NP2

8 y 9.

Los términos resultantes son: *el, 17, de, octubre*; mismos que no son relevantes. En el caso de la preposición en que precede a un nombre propio, se le puede considerar indicativo de lugar, por lo cual NP2 señala donde se llevó a efecto la acción.

Al final del procesamiento, se obtiene lo mostrado en la tabla 5.10.

Elemento	Subelemento	Contenido
Lugar		NA
Fecha		NA
Agente		Roberto_Adame_Doria,_director_de_ADC_Consultores_de_Negocios
Acción		señalar
Objeto		
	OR= Referente	La formación
	OD = Determinativo del ref.	de_un_emprendedor_eficiente_comienza_en_la_familia,_los_estudios_en_segundo_lugar_y_las_empresas_incubadoras_después
	OC = Calificación	NA
	ODC= Determinativo de la Cal.	NA
Paciente		
Complemento		
	I = Instrumento	NA
	E = Efecto	NA
	K = Contexto	NA
Relator		NA

Tabla 5.10 Elementos de una noticia de estructura O-V-A

5.9.6 EJEMPLO 3. NOTICIA DE TIPO VOA

Procedimiento.

1. Encontrar y englobar los nombres propios
2. Etiquetar y sustituir nombres propios
3. Detectar las locuciones no verbales y sustituir
4. Encontrar los verbos y sustituir
5. Determinar el objeto, en todos y cada uno de sus elementos
6. Determinar el verbo principal y los secundarios
7. Distinguir los nombres propios
8. Definir el lugar y fecha del evento
9. Analizar los términos no participantes

Texto original

“Mercadotecnia Estratégica”. Sobre esta temática versó el curso de Capacitación que impartió hace algunos días, el Dr. Román Esqueda Ataide, Director del despacho Posicionamiento y Consultores S.C. de la ciudad de México.

1. Se inicia englobando los términos indicativos de nombre propio.

“Mercadotecnia_Estratégica”. Sobre esta temática versó el curso de Capacitación que impartió hace algunos días, el Dr. Román Esqueda Ataide, Director del despacho Posicionamiento y Consultores S.C. de la ciudad de México.

2. Luego de la sustitución y etiquetado de nombres propios, resulta:

NP1 *Sobre esta temática versó* NP2 *que impartió hace algunos días*, NP3 *de*
NP4

NP1 = “Mercadotecnia_Estratégica”.

NP2 = el_curso_de_Capacitación

NP3 =

el_Dr._Román_Esqueda_Ataide,_Director_del_despacho_Posicionamiento_y_Consultores_S.C.

NP4 = la_ciudad_de_México.

3. Inicialmente se buscan en los nombres propios, algunos de los objetos referenciados en el anexo B. De no encontrarse, se busca en los términos restantes. En NP2 se detectan dos objetos (*curso* y *capacitación*), por lo que se procede a delimitar el texto para su procesamiento, resultando en:

OBJETO = NP2 = el_curso_de_Capacitación

Del anterior objeto, se buscan los elementos previos y conectores de sus componentes: referente, determinativo y calificativo, resultando:

OBJETO REFERENTE (OR)

Previo	[NC]	extensor	[NC]
<i>El</i>	<i>curso</i>	<i>de</i>	<i>Capacitación</i>

Enseguida se buscan aquellos términos que conectan al referente con el determinativo, encontrándose en este caso la palabra *que* y resultando en:

OBJETO DETERMINATIVO (OD)

previo	[TEXTO]
<i>que</i>	<i>impartió hace algunos días,</i>

NP1 *Sobre esta temática* versó OBJETO[OR,OD], NP3 *de* NP4

OR = el_curso_de_Capacitación

OD = impartió_hace_algunos_días

4 y 5. Se buscan las locuciones no verbales y al no encontrarlas, se procede a la detección, etiquetado y sustitución de los verbos, quedando:

NP1 *Sobre esta temática* V1 OBJETO[OR,OD], NP3 *de* NP4

V1 = versar

6. Al tenerse un solo verbo se le considera como el principal, por lo cual

NP1 *Sobre esta temática* VERBO OBJETO[OR,OD], NP3 *de* NP4

VERBO = V1

7. Para distinguir los nombres propios, se procesan NP1, NP3 y NP4. NP1 al estar acotado entre comillas, se considera solamente un tema o contenido.

TEMA *Sobre esta temática* VERBO OBJETO[OR,OD], NP3 *de* NP4

TEMA = NP1

Al no existir locuciones, del texto se buscan las preposiciones, conjunciones y adverbios que aparecen como elementos previos (Anexo E) y se procede a la asignación.

La preposición *sobre* no conecta con nombre propio por lo cual no sirve para distinguir los nombres propios.

La preposición *de* puede implicar pertenencia u origen, por lo cual NP3 puede absorber a NP4. Para este caso, la estructura de la noticia corresponde a V-O-A, por lo cual NP3 y NP4 representan el agente.

TEMA *Sobre esta temática* VERBO OBJETO[OR,OD], AGENTE

AGENTE =

el_Dr._Román_Esqueda_Ataide,_Director_del_despacho_Posicionamiento_y
_Consultores_S.C. de la_ciudad_de_México.

El verbo principal afecta solo al intelecto, por lo cual se considera a la noticia como un hecho intangible. En éste, es posible encontrar un relator por lo cual, se busca entre los términos restantes (*Sobre, esta y temática*) y al no encontrarse toma un valor nulo.

Es importante considerar que cuando en la noticia se refiere un tema, se considera a éste como el determinativo del objeto, sobre cualquier otro segmento del texto, por lo cual al final resulta.

OBJETO[OD] *Sobre esta temática* VERBO OBJETO[OR], AGENTE

8. Se buscan indicativos de lugar y fecha del evento, pero al no encontrarse se declaran nulos.

9. Los términos restantes (Sobre esta temática) son insignificantes para el trabajo.

Al final del procesamiento, las etiquetas se sustituyen por su contenido y se obtienen los elementos básicos del hecho noticioso planteado inicialmente, como se muestra en la tabla 5.11.

Elemento	Subelemento	Contenido
Lugar		NA
Fecha		NA
Agente		el_Dr._Román_Esqueda_Ataide,_Director_del_despacho_Posicionamiento_y_Consultores_S.C._de la_ciudad_de_México.
Acción		versar
Objeto		
	OR= Referente	El_curso_de_Capacitación
	OD = Determinativo del ref.	“Mercadotecnia_Estratégica”.
	OC = Calificación	NA
	ODC= Determinativo de la Cal.	NA
Paciente		NA
Complemento		NA
	I = Instrumento	NA
	E = Efecto	NA
	K = Contexto	NA
Relator		NA

Tabla 5.11 Elementos de una noticia de estructura V-O-A

5.9.7 EJEMPLO 4. NOTICIA DE TIPO VAO

Procedimiento.

1. Encontrar y englobar los nombres propios
2. Etiquetar y sustituir nombres propios
3. Determinar el objeto, en todos y cada uno de sus elementos
4. Detectar las locuciones no verbales y sustituir
5. Encontrar los verbos y sustituir
6. Determinar el verbo principal y los secundarios
7. Distinguir los nombres propios
8. Definir el lugar y fecha del evento
9. Analizar los términos no participantes

Texto original

"El Budismo ha llegado a Occidente para quedarse" afirmó el maestro coreano Samu Sunim, en conferencia frente a estudiantes y catedráticos de la Universidad de las Américas Puebla (UDLA).

1.

"El_Budismo_ha_llegado_a_Occidente_para_quedarse" afirmó el maestro coreano Samu_Sunim, en conferencia frente a estudiantes_y_catedráticos_de_la_Universidad_de_las_Américas_Puebla_(UDLA).

2.

NP1 *afirmó el maestro coreano* NP2, *en conferencia frente a* NP3

NP1 = *"El_Budismo_ha_llegado_a_Occidente_para_quedarse"*

NP2 = *Samu_Sunim*

NP3 =

estudiantes_y_catedráticos_de_la_Universidad_de_las_Américas_Puebla_(UDLA).

3.

Se busca un término que represente un objeto, entre los nombres propios y de no encontrarse, entre los demás términos, de lo que se desprende que:

OR = conferencia

NP1 *afirmó el maestro coreano* NP2, en OBJETO[OR] *frente a* NP3

4. De entre los términos restantes, se buscan las locuciones y se sustituyen

NP1 *afirmó el maestro coreano* NP2, en OBJETO[OR] L1 NP3

L1 = *frente_a*

5.

Se detectan los verbos y se sustituyen por sendas etiquetas.

NP1 V1 *el maestro coreano* NP2, en OBJETO[OR] L1 NP3

V1 = afirmar

6.

Al tener un solo verbo, se consigna como verbo principal

NP1 VERBO *el maestro coreano* NP2, en OBJETO[OR] L1 NP3

VERBO = V1

7.

Para distinguir los nombres propios, se procesan NP1, NP2 y NP3. El primero al estar acotado entre comillas indica una alta probabilidad de tratarse del tema (determinativo del objeto).

OBJETO[OD] VERBO *el maestro coreano* NP2, en OBJETO[OR] L1 NP3

OD = NP1

Enseguida, las locuciones no verbales encontradas se buscan en los términos previos a elementos básicos (Anexo E). Para este caso, *frente_a* se encuentra como previo de un paciente, por tanto resulta:

OBJETO[OD] VERBO *el maestro coreano* NP2, en OBJETO[OR] L1
PACIENTE

PACIENTE = NP1

Finalmente, del texto se buscan las preposiciones, conjunciones y adverbios que aparecen como elementos previos (Anexo E) y se procede a la asignación. Para este caso, el único elemento es *en*, el cual aparece como previo a un objeto o paciente. En el texto, se encuentra como previo de un objeto, por lo que se descarta que vaya antes de un paciente.

El nombre propio restante NP3 se califica como el agente de la noticia, porque en las estructuras de tipo V-A-O, el agente es el nombre propio encontrado luego del verbo.

OBJETO[OD] VERBO *el maestro coreano* AGENTE, en OBJETO[OR] L1
PACIENTE

AGENTE = NP2

Al final del procesamiento, se obtiene lo mostrado en la tabla 5.12

Elemento	Subelemento	Contenido
Lugar		NA
Fecha		NA
Agente		Samu_Sunim
Acción		afirmar
Objeto		
	OR= Referente	conferencia
	OD = Determinativo del ref.	"El_Budismo_ha_llegado_a_Occidente_para_quedarse"
	OC = Calificación	NA
	ODC= Determinativo de la Cal.	NA
Paciente		estudiantes_y_catedráticos_de_la_Universidad_de_las_Américas Puebla (UDLA).
Complemento		
	I = Instrumento	NA
	E = Efecto	NA
	K = Contexto	NA
Relator		NA

Tabla 5.12 Elementos de una noticia de estructura V-A-O

5.10 MANEJO DE RESULTADOS

5.10.1 NORMALIZACIÓN DE LOS RESULTADOS

Debido a que el contenido de algunos elementos básicos (Verbos y Referentes de objeto), se puede considerar atómico y repetible, es necesario normalizarlos en base a los siguientes criterios.

Verbos.

Como se ha mencionado previamente, el verbo se trabaja en infinitivo y las locuciones verbales se sustituyen con sinónimos, tal como se documenta en el anexo G.

Ejemplo.

El verbo: *presentará*

Resulta en: *presentar*

La locución verbal: *dar a conocer*

Resulta en: *informar*

Referentes de objeto.

Aun cuando el referente hasta este momento ha conservado el artículo que le precede. Se procede a la eliminación de cualquier término que no aparezca en el anexo de objetos y se deja el término en singular.

Ejemplo.

El referente: *la Celebración Eucarística*

resulta en: *celebración*

Calificativos de objeto.

Los adjetivos se normalizan mediante su lematización. El lema de los adjetivos (calificativos y ordinales) es siempre la forma masculina singular o la forma singular si es un adjetivo de una sola terminación.

Ej.

El adjetivo: *nuevas*

Resulta en: *nuevo*

5.10.2 REGISTRO DE LAS NOTICIAS

Una vez normalizados los elementos básicos obtenidos, se procede al registro de los siguientes atributos, tal y como se muestra en la figura 5.11:

Empresa	= La entidad que publica la noticia
Fecha	= La fecha de la publicación
Noticia	= Consecutivo de noticia por empresa y fecha
Agente	= El descriptivo de quien realizó la acción
Verbo	= La actividad realizada por el agente
Objeto Referente	= El tipo de objeto usado
Objeto Determinativo	= La descripción amplia del objeto
Paciente	= El descriptivo de quien recibe el efecto de la acción

Para un mayor abundamiento, en el anexo J, se plasman los resultados de procesar un total de 130 noticias del 1 de marzo de 2006, para las empresas antes citadas.

empresa	fecha	agente	verbo	objeto ref.
UCC	01/03/2006	La Dirección General de Extensión Universitaria a través del Centro de	haber, preparar	deportes y actividades recreativas
UCC	01/03/2006	La Universidad Cristóbal Colón	presentar	la convocatoria semestral
UCC	01/03/2006	La Parroquia Universitaria	invitar	las celebraciones eucarísticas
UCC	01/03/2006	El rey Carlos III	encomendar	la traducción
UCC	01/03/2006		Acreditar	Certificación Microsoft Office Specialist nivel

objeto det.	paciente
participar en programas de movilidad	los alumnos de licenciatura
la Biblia completa al castellano	
Com	

Figura 5.11. Registro de los resultados

Sin embargo y como puede observarse, algunos de ellos tales como Empresa, Verbo y Referente de Objeto, por su naturaleza, presentan duplicidad en su contenido, por lo cual pueden ser referenciados mediante claves en otros archivos, tal y como se aprecia en la figura 5.12 y cuyo mayor detalle se puede encontrar en el anexo K, donde aparecen las sustituciones realizadas.

empresa	fecha	noticia	agente	verbo	objeto ref.
2	01/03/2006	1	La Dirección General de Extensión Universitaria a través del Centr	111	54
2	01/03/2006	2	La Universidad Cristóbal Colón	67	48
2	01/03/2006	3	La Parroquia Universitaria	52	26
2	01/03/2006	4	El rey Carlos III	98	151

objeto det.	paciente
participar en programas de movilidad	los alumnos de licenciatura

Figura 5.12. Registro de los resultados con sustitución

Adicionalmente, a partir del análisis de los resultados se observa que los componentes del contexto, solamente sirven para expresar más ampliamente la noticia, pero no representan elementos sustanciales para la función de inteligencia propuesta, por lo cual se toma la decisión de no incluirlos en el registro de la base de datos.

Para el registro de cada noticia, se realicen las siguientes comandos:

/ Se incrementa la clave de noticia por empresa */*

Noticia = Noticia + 1

/ La fecha es del día en que se publicó la noticia */*

insert into NOTICIAS values (:ClaveEmp, :Fecha, :Noticia, :Agente, :ClaveVerbo, :ClaveObjRef, :ObjDet, :ObjCal, :Paciente)

5.11 EXPLOTACIÓN DE LOS DATOS

5.11.1 PLANTEAMIENTO DE PREGUNTAS GENÉRICAS DE INTELIGENCIA COMPETITIVA

De acuerdo con Castresana [47], existen ciertos momentos que vive la empresa que son idóneos para la difusión de una noticia. Estos corresponden a eventos que desde la inteligencia competitiva, sirven de guía para el planteamiento de preguntas generales que ha de responder el sistema que se propone. Enseguida se muestra la clasificación de los eventos así como la pregunta genérica que propician. Por supuesto, que la manera definitiva de clasificarles y las preguntas a responder, estarán supeditadas al criterio y necesidades del administrador del sistema.

Pregunta

- a. ¿En que actividades está involucrada, como participante u organizadora?

Tipos de evento

- Participación u organización en actos culturales.
 - Participación u organización en actos sociales.
 - Participación u organización en actos benéficos.
 - Participación u organización en actos deportivos.
 - Participación en ferias, congresos o conferencias.
 - Participación en seminarios o conferencias.
 - Realización de jornadas de puertas abiertas.
-
- b. ¿Qué desea difundir como novedad?
 - Lanzamiento de un nuevo producto o servicio
 - Apertura de nuevas filiales.
 - Inversiones en nuevos negocios.
 - Ampliaciones de personal
 - Nuevo nombre o imagen corporativa.
 - Ampliaciones de capital.
 - Mejoras técnicas, de calidad o de organización en su servicio.
 - Expansión a nuevos mercados o líneas de negocio.
 - Compra de empresas de la competencia.
 - Contrataciones y ceses de directivos.
 - Alianzas empresariales.
 - Implementación de un programa de innovación.

- c. ¿Qué distinciones se han obtenido?
- Consecución de nuevos clientes de reconocido prestigio.
 - Obtención de premios o reconocimientos.
 - Obtención de certificados de calidad (ISO, AENOR, etc.)
 - Consecución de un récord de cualquier tipo.
 - Obtención de premios, nominaciones y reconocimientos.
 - Evolución positiva de su cifra de ventas.
- d. ¿Qué se celebra?
- Aniversarios de la creación de la propia empresa.
 - Aniversarios de productos o servicios.
- e. ¿De qué manera beneficia a la sociedad?
- Consecución de mejoras para la sociedad.
 - Donaciones.
 - Patrocinios.
 - Acciones ecológicas o encaminadas a mejorar la seguridad.
 - Firma convenios de colaboración.
 - Realización de estudios de impacto social.
- f. ¿Cuáles han sido los resultados obtenidos de investigaciones?
- Resultados de investigaciones científicas.
 - Resultados sobre estudios de mercado.
 - Resultados de juntas de accionistas.
- g. ¿Qué se está haciendo para mejorar la imagen y opinión pública de la empresa?
- Convocatorias de prensa.
 - Informe sobre resultados económicos.
 - Declaraciones destacadas de directivos.
 - Desmentido de rumores y calumnias.
 - Comentarios ante problemas coyunturales de su sector.
 - Comentario de sentencias judiciales
 - Publicación de anuncios oficiales.
 - Anuncio de futura junta de accionistas.
 - Publicación de resoluciones judiciales.
- h. ¿Qué se está haciendo para fines de promoción?
- Presentación de una campaña de publicidad o marketing.
 - Descuentos y rebajas.
 - Presentaciones comerciales.
 - Previsiones de facturación.

5.11.2 CRITERIOS Y COMANDOS PARA DAR RESPUESTA

Para la explotación de la base de datos y la generación de respuestas a las preguntas previamente planteadas, se propone lo siguiente:

Los eventos que aparecen registrados para cada pregunta genérica se pueden obtener a partir de consultas en SQL al archivo de noticias, a partir de los criterios que, a manera de ejemplo, se muestran enseguida. Para una mejor explicación, se presentan los comandos, considerando los elementos descriptivos en lugar de sus respectivas claves.

a. ¿En que actividades está involucrada, como participante u organizadora?

Criterios.

Objeto Referente = Aquellos que implican un evento de tipo cultural, social, cultural, científico, tecnológico, deportivo.

Verbo = Aquellas acciones que impliquen la organización o participación

Consulta.

*Select **

From Noticias

Where ObjetoRef in ('conferencia', 'congreso', 'encuentro', 'partido',...)

And verbo in ('organizar', 'celebrar', 'participar', ...)

b. ¿Qué desea difundir como novedad?

Criterios.

Objeto Calificativo = Aquellos términos que impliquen la implementación de un objeto referente novedoso o la ampliación y mejora de uno existente

Verbo = Aquellas acciones que impliquen la implementación de una entidad nueva

Consulta.

*Select **

From Noticias

Where ObjetoCal in ('nuevo', 'innovador', 'cambio', 'ampliación', 'apertura', 'inicio'...)

And verbo in ('crear', 'iniciar', 'implementar', 'desarrollar' ...)

c. ¿Qué distinciones se han obtenido?

Criterios.

Objeto Referente = Aquellas palabras que determinen algún tipo de logro

Verbo = Aquellas acciones que impliquen la consecución de un objetivo

Consulta.

*Select **

From Noticias

Where ObjetoRef in ('premio', 'medalla', 'reconocimiento', 'campeonato', 'certificado', 'victoria'...)

And verbo in ('lograr', 'obtener', 'conseguir', 'alcanzar' ...)

d. ¿Qué se celebra?

Criterios.

Objeto Referente = Los términos que determinen el alcance de un objetivo en función del tiempo

Verbo = Aquellas acciones que impliquen la celebración de un periodo más de vida

Consulta.

*Select **

From Noticias

Where ObjetoRef in ('aniversario', 'celebración', 'festividad', 'fiesta', 'año', 'lustró'...)

And verbo in ('celebrar', 'conseguir', 'alcanzar', ...)

e. ¿De qué manera se beneficia a la sociedad?

Criterios.

Objeto Referente = Aquellas palabras que determinen beneficios a grupos sociales

Verbo = Aquellas acciones que impliquen actividades de beneficio

Consulta.

*Select **

From Noticias

Where ObjetoRef in ('donacion', 'convenio', 'estudio', 'patrocinio', 'beneficencia', 'apoyo'...)

Or verbo in ('apoyar', 'firmar', 'donar', 'patrocinar', ...)

f. ¿Cuáles han sido los resultados obtenidos de investigaciones?

Criterios.

Objeto Referente = Los términos que den a entender resultados de un proceso investigativo

Verbo = Las acciones que determinen el dar a conocer una meta

Consulta.

*Select **

From Noticias

Where ObjetoRef in ('resultado', 'logro',...)

and verbo in ('presentar', 'informar', 'dar a conocer', 'patrocinar', ...)

g. ¿Qué se está haciendo para mejorar la imagen y opinión pública de la empresa?

Criterios.

Objeto Referente = Los términos que impliquen eventos cuya finalidad sea mejorar el nombre de la empresa

Verbo = Las acciones que den a entender la difusión del evento asociado al objeto referente

Consulta.

*Select **

From Noticias

Where ObjetoRef in ('declaracion',

'desmentido', 'comentario', 'publicacion',...)

and verbo in ('presentar', 'informar', 'dar a conocer', ...)

h. ¿Qué se está haciendo para fines de promoción?

Criterios.

Objeto Referente = Los términos que impliquen eventos cuya finalidad sea mejorar el nombre de la empresa

Verbo = Las acciones que den a entender la difusión del evento asociado al objeto referente

Consulta.

*Select **

From Noticias

Where ObjetoRef in ('campana', 'descuento', 'rebaja',...)

and verbo in ('presentar', 'informar', 'dar a conocer', ...)

Los resultados se pueden restringir aún más, agregando otros criterios tales como la(s) empresa(s) de interés y el rango de fechas deseadas. Para ello, basta con agregar a los anteriores comandos, alguna o ambas de las siguientes condiciones:

And Empresa in ('Lista de empresas de interés')

And fecha between (Fecha Inicial) and (Fecha de Final)

Los resultados de las consultas así construidas aparecen de manera similar a la mostrada en la figura 5.13.

The screenshot shows a web browser window titled "Consulta de noticias - Microsoft Internet Explorer". The address bar shows "http://1127.0.0.1/ucc/index.php". The page content includes the UCC logo and the title "Sistema de consulta de noticias". Below the logo is a search form with the following fields:

- Selección de pregunta: ¿En que actividades está involucrada, como participante u organizadora?
- Institución: (Todas)
- Del (yyyy-mm-dd): 2006-03-01
- Hasta (yyyy-mm-dd): 2006-03-01

Below the form is a button labeled "Iniciar búsqueda". The results are displayed in a table with the following data:

ID	Fecha	Institución	Agente	Verbo	Referente	Detalle	Paciente
4	2006-03-01	UCLA	El Cine Club de la UCLA	proyectar	película	"Perdón María"	
6	2006-03-01	UCLA	La Compañía de Teatro de la Universidad de las Américas Puc	presentar	obra	"Las Preciosas Ridículas"	
2	2006-03-01	EBC	la EBC	realizar	convivencia	los alumnos	Bachillerato y Licenciatura
3	2006-03-01	UPAZP	La artista Pepita Albino	donar	obra		el museo

Fig. 5.13 Ejemplo de resultados de las consultas

5.11.3 CONSTRUCCIÓN DINÁMICA DE LOS COMANDOS

Se propone, en primer lugar, que la introducción y actualización de los elementos que conforman cada uno de los criterios la efectúe el usuario administrador del sistema. Para ello, se necesita registrar:

- Los verbos que se han de verificar por cada pregunta
- Los referentes que se necesitan discriminar en cada cuestionamiento

en tablas como BASADAS_EN y BASADAS_EN1 cuyas estructuras y vinculaciones con otras, se observa en la figura 5.14.

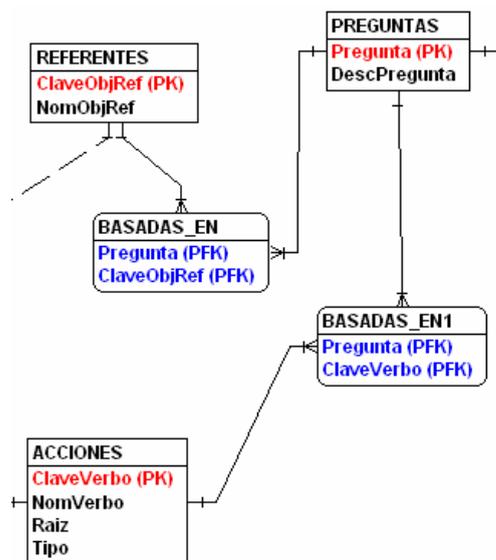


Figura 5.14 Estructura y vinculaciones de BASADAS_EN y BASADAS_EN1

En segundo término, se propone la construcción dinámica de los comandos SQL anteriormente citados. A continuación, se muestra en la figura 5.15 los pasos a seguir y se complementa con un ejemplo.

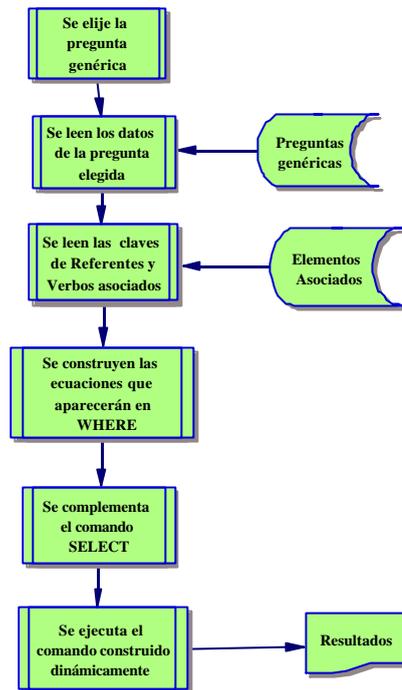


Figura 5.15 Pasos para la consulta

Ejemplo.

Pregunta a. ¿En que actividades está involucrada, como participante u organizadora?

Criterios.

Objeto Referente = Aquellos que implican un evento de tipo cultural, social, cultural, científico, tecnológico, deportivo.

Verbo = Aquellas acciones que impliquen la organización o participación

Consulta.

*Select **

From Noticias

Where ObjetoRef in ('conferencia', 'congreso', 'encuentro', 'partido',...)

And verbo in ('organizar', 'celebrar', 'participar', ...)

Para adecuar el anterior comando, primero se requiere cargar en la tabla PREGUNTAS, el contenido de la pregunta.

```
insert into PREGUNTAS values (1, '¿En que actividades está involucrada, como participante u organizadora?')
```

Luego y a partir del contenido de las tablas ACCIONES y REFERENTES, guardar y/o actualizar el contenido de las tablas BASADAS_EN y BASADAS_EN1, con los comandos:

```
/* Para guardar */
insert into BASADAS_EN values (1,:Claveobjref)
insert into BASADAS_EN1 values (1, :ClaveVerbo)

/* Para actualizar */
update BASADAS_EN set ClaveobjRef= :Claveobjref where Pregunta = :pregunta
update BASADAS_EN1 set ClaveVerbo= :ClaveVerbo where Pregunta = :pregunta
```

De modo que el contenido de ambos archivos, podría ser como el mostrado en la tabla 5.13.

Pregunta	ClaveObjRef	Pregunta	ClaveVerbo
1	1	1	29
1	65	1	16
1	12	1	23
1	17	1	14
1	18	1	11

Tabla 5.13 Contenido de BASADAS_EN y BASADAS_EN1

Para finalmente, leyendo cíclicamente el contenido de ambas tablas,

```
Select ClaveObj
From BASADAS_EN
Where Pregunta = 1
into :ClaveObjRef

Select ClaveVerbo
From BASADAS_EN1
Where Pregunta = 1
into :ClaveVerbo
```

construir las condiciones de forma dinámica, así:

```
Var
  Objetos:string;
  Verbos:string;
Begin
  Objetos:= Objetos + ClaveObjRef + ' ';
  Verbos:= Verbos + ClaveVerbo + ' ';
```

Resultando el comando a ejecutar en:

```
Var
  Vsql:string;
Begin
  Vsql :=
    'Select N.Fecha, E.NombreEmp, N.Noticia, N.Agente, A.NomVerbo,
    R.NomObjRef, N.ObjDet, N.ObjCal, N.Paciente
    From EMPRESAS E, A.ACCIONES, REFERENTES R, NOTICIAS N
    Where E.ClaveEmp = N.ClaveEmp
    And A.ClaveVerbo= N.ClaveVerbo
    And R.ClaveObjRef = N.ClaveObjRef
    And ObjetoRef in ('+ objetos +')';
  Vsql := vsql + ' And verbo in ('+ verbos +')';
  Query.close;
  Query.sql.text := vsql;
  Query.open;
End;
```

Para el caso de mostrar solamente ciertas empresas, se puede implementar la condición correspondiente de manera análoga.

```
Empresas = '(:Empresa1, :Empresa2, :Empresa3, :Empresa4...)'
```

Y agregarla luego de los verbos, en el comando previo, de la siguiente manera:

```
'and Empresa in ('+ Empresas +)';
```

5.12 ARQUITECTURA GENERAL DEL SISTEMA.

El sistema está conformado de tres módulos: configuración, procesamiento y consultas. Los dos primeros, a cargo del usuario administrador y el último para difundir los hallazgos de inteligencia a los ejecutivos de la empresa. Las particularidades de cada módulo se describen enseguida:

5.12.1 MÓDULO DE CONFIGURACIÓN

El usuario será capaz de mantener al día los archivos de referencia (catálogos) mediante una aplicación en lenguaje Delphi V6.0 apoyado con el cliente de base de datos Firebird V1.5 en arquitectura cliente-servidor, pues mientras el *front-end* se ejecutará en la computadora personal del administrador (Windows-Intel), el *back-end* correrá en un equipo anfitrión con software de base Linux Debian 3.1.3. En la figura 5.16, se muestra como interactúa el módulo con el usuario y con la base de datos.

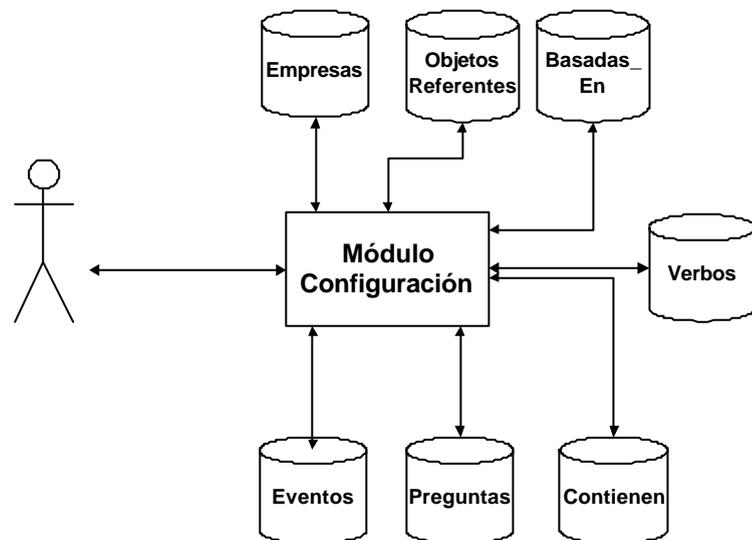


Fig. 5.16. Interacciones en datos del módulo de Configuración

Como puede apreciarse, el usuario a través de éste módulo, consulta, registra y actualiza el contenido de prácticamente todos los archivos usados como referencia.

El módulo de configuración maneja las interfaces de usuario gráficas que se presentan en la siguiente sección. En el menú principal, aparecen dos opciones principales: CATÁLOGOS que permite el ABCD de los archivos mostrados y ACTUALIZACIÓN que permite exportar el contenido de los catálogos a archivos de texto plano, mismos que se usan en el módulo de procesamiento.

5.12.2 MÓDULO DE PROCESAMIENTO

Está compuesto de dos submódulos: lector y escritor. En una base diaria, se ejecutará el submódulo lector, el cual se encargará de conectarse a la página Web donde se publican las noticias de la empresa, hará una copia local del archivo correspondiente, lo transformará en un archivo texto sin etiquetas HTML. Finalmente, hará las comparaciones necesarias para mantener solamente los archivos y noticias, que no se hayan procesado ya. Este submódulo ha de implementarse en la plataforma de desarrollo Java 2.0, el cual genera dentro de la carpeta principal NOTICIAS, un directorio por cada grupo de noticias leídas por fecha y en donde se almacenarán los archivos correspondientes a cada empresa, así:

```
NOTICIAS-> 20061008    -> UPAEP.TXT
                        -> UCC.TXT
                        -> UDLA.TXT
                        etc.
```

Los intercambios de datos que efectúa el submódulo Lector aparecen en la figura 5.17. Por cada registro del archivo Empresas, obtiene las direcciones de los sitios electrónicos donde cada empresa publica sus noticias y realiza una copia de la página Web en formato de texto.

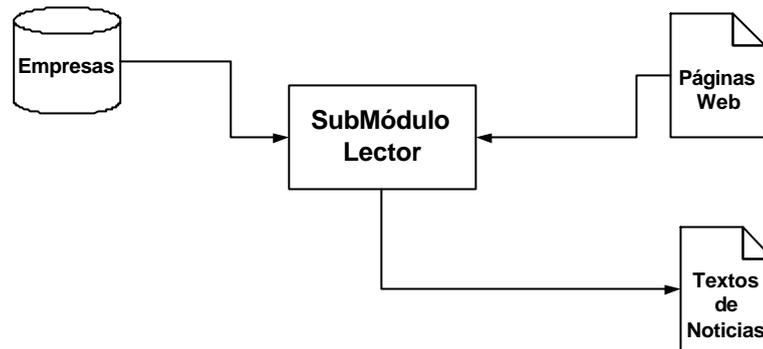


Figura 5.17 Interacciones en datos del submódulo Lector

El submódulo escritor (figura 5.18) se apoya en los archivos exportados de la base de datos mantenida por el módulo de Configuración. A partir de los Referentes de objetos, Empresas y Acciones, se procesa el titular y/o encabezamiento de la noticia con el algoritmo denominado “Máxima Abstracción”. Luego se normalizan ciertos términos, se buscan las claves que corresponden a los descriptivos de los términos normalizados y se registran los elementos en la tabla Noticias. El submódulo de referencia ha de construirse mediante la plataforma Regina 3.3, cuyo lenguaje interpretado REXX tiene grandes facilidades de manipulación de textos.

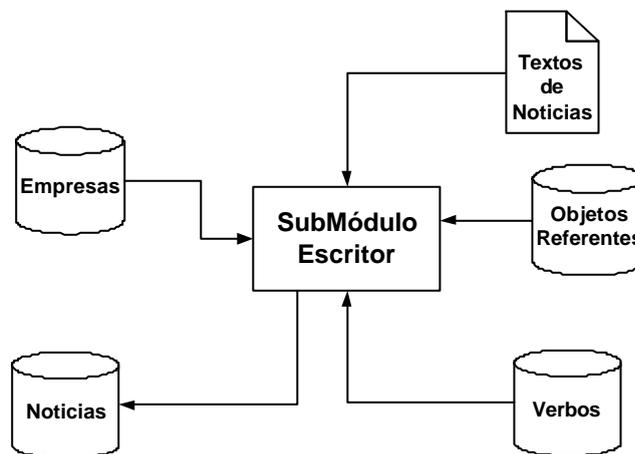


Figura 5.18 Interacciones en datos del submódulo Escritor

Como puede apreciarse, el citado submódulo se apoya en los archivos de Verbos y Objetos referentes, para luego del procesamiento registrar de manera estructurada, las noticias por cada empresa.

5.12.3 MÓDULO DE CONSULTAS

Este módulo comprende un conjunto de consultas que permitirá dar respuesta a las preguntas propuestas de Inteligencia Competitiva. Las interfaces de usuario serán accedidas por los usuarios ejecutivos, luego de validar sus claves de usuario y contraseña. Los usuarios pueden luego de elegir la pregunta, filtrar aún más los resultados, introduciendo la empresa de interés y un rango de fechas determinado.

El módulo se apoya en la actualización que el administrador hace de las tablas Empresas, Verbos, Objetos Referentes, Preguntas, Eventos, pero sobre todo, del mantenimiento a las asociaciones Pregunta-Verbo, Pregunta-Referente y Pregunta-Evento. Conforme, los ejecutivos lo requieren, el administrador puede agregar, modificar ó eliminar preguntas y sus criterios para la generación de respuestas. Este módulo se desarrollará en lenguaje PHP V4.0 con la incrustación de comandos SQL contruidos de manera dinámica, los cuales se envían al servidor de base de datos FireBird V1.5 y cuyas respuestas permiten construir páginas Web dinámicas, a ser consultadas por los ejecutivos.

En la figura 5.19, se puede apreciar la manera como el módulo de consultas lee, a partir de las claves de cada noticia, las descripciones de empresas, verbos y objetos referentes correspondientes. La decisión de filtrar cada noticia se da con base a las claves de verbos y referentes asociados a la pregunta genérica que el usuario ejecutivo elige hacer al sistema, los cuales fueron vinculados en el módulo de configuración, por el administrador. Las noticias pueden filtrarse aún más a partir de un periodo determinado por el ejecutivo.

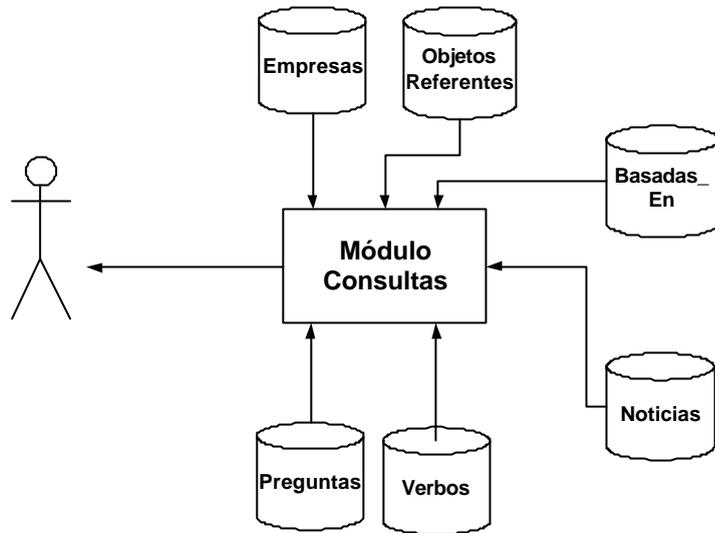


Figura 5.19 Intercambio de datos del módulo de Consultas

5.12.4 PLATAFORMA DE EJECUCIÓN

La plataforma (*Hardware, software & networking*) que se propone para implementar la arquitectura cliente/servidor se muestra en la figura 5.20, en la cual las peticiones de los diferentes clientes, son atendidas por los servidores de:

- | | |
|--|---------------|
| - Base de datos | Firebird V1.5 |
| - Protocolo de transferencia de hipertexto | Apache V2.0 |
| - Transacciones | Php V4.0 |

mismos que son alojadas en el equipo anfitrión (HOST) de la figura.

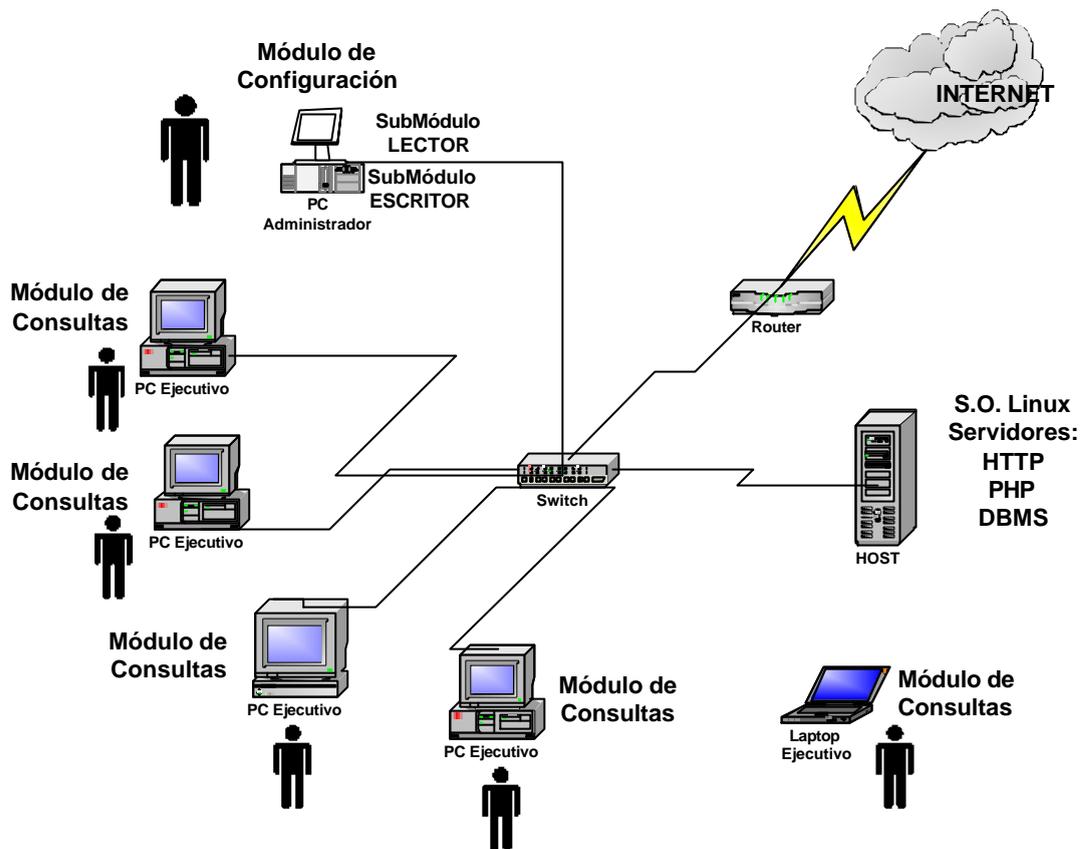


Figura 5.20 Plataforma para la ejecución del sistema

Las aplicaciones cliente que corresponden a los módulos de configuración y procesamiento, son ejecutadas en el equipo asignado al administrador del sistema propuesto; mientras que el programa cliente para la consulta de los ejecutivos se obtiene a partir de la conexión al servidor Web (HOST).

La conexión a Internet mostrada, se requiere para la operación diaria de obtener una copia local de la página Web de cada empresa en particular.

5.13 INTERFACES DE USUARIO DEL SISTEMA.

El sistema es accedido por el usuario administrador para realizar tareas de mantenimiento de catálogos, actualización de archivos de textos y respaldo y restauración de la base de datos. Para ello se utiliza el menú principal como puede verse en la figura 5.21

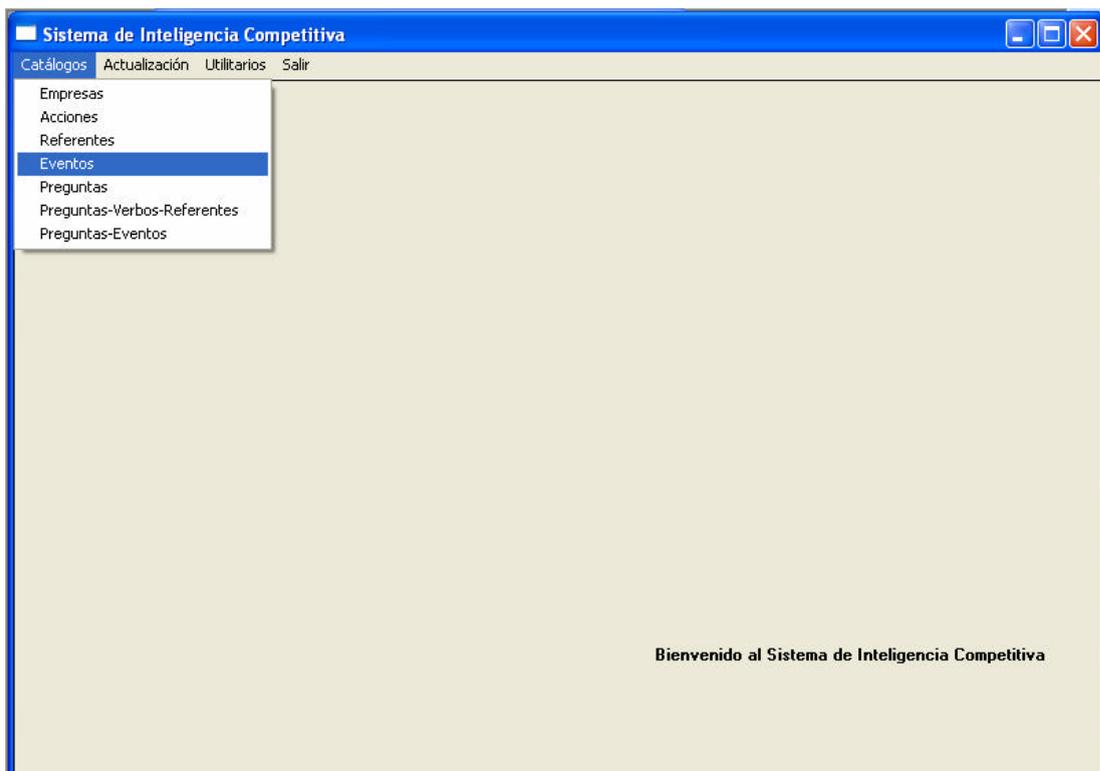


Figura 5.21 Menú principal del sistema

Cuando el usuario requiera llevar a cabo el mantenimiento de los catálogos, aparecerán interfaces gráficas de usuario como las mostradas en la figura 5.22.

En el momento que se desee asociar a una pregunta una serie de verbos y referentes de objetos, se ingresará a una pantalla como la mostrada en la figura 5.23.

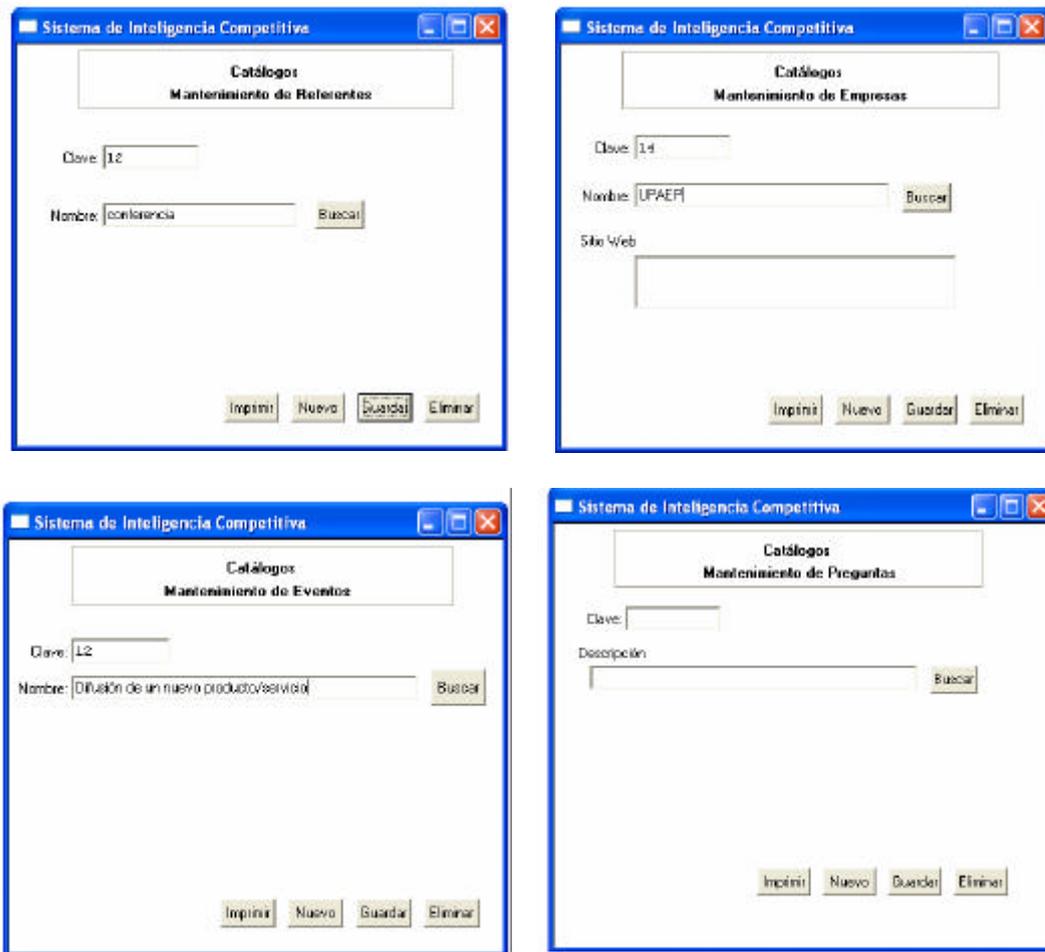


Fig. 5.22. Interfaces de usuario de Catálogos

De la misma manera, cuando se desee asociar múltiples eventos a una pregunta, se podrá realizar mediante la interfaz de usuario mostrada en la figura 5.24.

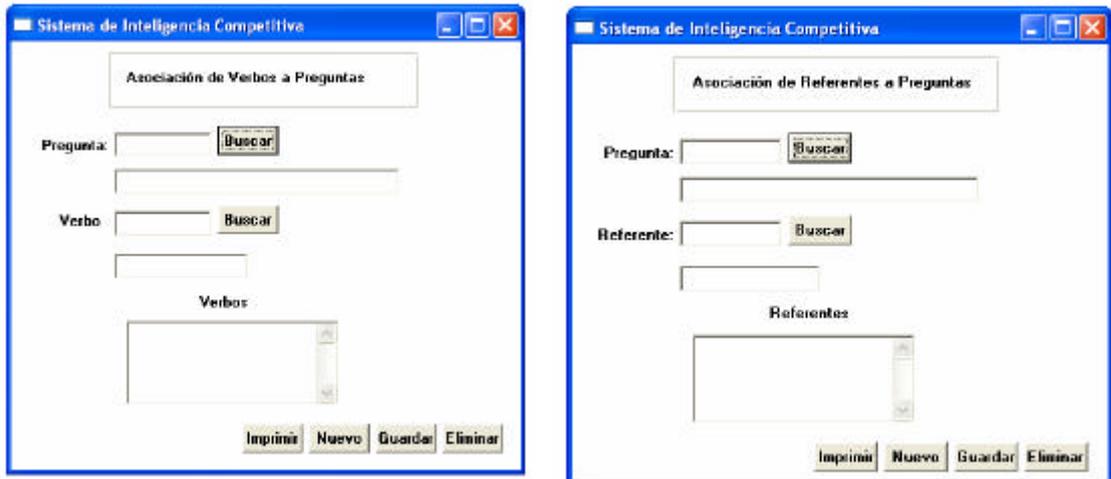


Figura 5.23 Pantallas para asociar verbos y referentes a preguntas



Figura 5.24 Interfaz de usuario para Pregunta- Eventos

Capítulo VI.

Conclusiones, recomendaciones y trabajo futuro.

En este capítulo se hace un recuento de las principales aportaciones hechas en este trabajo, se establecen las recomendaciones generales para la posible aplicación a otro tipo de empresas, se plantean algunas tareas futuras y se describe el impacto social que puede producir el trabajo

6.1 Conclusiones

Debido a las complicaciones que entraña el procesamiento de textos en lenguaje natural, la automatización actualmente, es posible sólo para lenguaje natural restringido.

Las noticias empresariales representan un dominio específico de conocimiento, que permite difundir la serie de acontecimientos (periódicos o eventuales) que se realizan en beneficio de sus consumidores y del entorno social en el cual se desarrolla.

Luego de un análisis profundo de la estructura de una noticia, de la naturaleza de sus componentes y de sus interrelaciones con términos de uso repetido, se determinó desarrollar un método orientado más hacia la sintaxis que a la ontología.

A continuación se detalla a manera de conclusiones, los siguientes apartados:

Respecto a la aplicación del trabajo.

Los resultados obtenidos del procesamiento automatizado mediante software construido en REXX, el cual dada su naturaleza de lenguaje interpretado y su característica manipulación de textos mediante un conjunto amplio de funciones, permiten concluir que la aplicación de la propuesta a otros ámbitos empresariales con publicaciones en el idioma español, es susceptible de efectuarse, tomando las providencias de cada caso. Adicionalmente, el diseño del algoritmo que se propone, pudo llevarse a la práctica, gracias a las prestaciones que ofrece la plataforma de desarrollo.

Respecto a los algoritmos generados.

De la muestra de noticias procesadas (131), el algoritmo propuesto permitió identificar el referente de objeto en un 85.5% y en un 96.2% el verbo principal.

Por tanto, se puede concluir que la efectividad de la propuesta está en función de:

- La exactitud en el procesamiento del método minero (Máxima Abstracción)
- La actualización permanente de los referentes de objetos
- El enriquecimiento del conjunto de verbos y formas verbales
- Y principalmente, la vigilancia permanente para identificar cambios en la dirección electrónica donde la empresa ubica sus noticias empresariales.

Para ello, el método propuesto tiene integrados indicadores de falla en la detección del portal de noticias, de los referentes y verbos, así como de una falla completa de la transformación del texto a una forma estructurada.

Es de notarse que en un reducido número de noticias, fue muy difícil para el autor, entender el mensaje que se deseaba comunicar, en cuyo caso disminuirá, por razones naturales, la efectividad del método.

Respecto al producto final.

La minería de textos con el método que se propone, se ve enriquecido con los módulos que permiten el mantenimiento a los datos de referencia y sobre todo, con el correspondiente a, el planteamiento de preguntas y la generación de respuestas, lo cual permitirá a los ejecutivos tener a su alcance, información actualizada de lo que acontece en el sector de negocios de su incumbencia.

Adicionalmente, se concluye que para transformar esta propuesta en una solución de llave en mano, susceptible de comercializarse, es de vital importancia la designación, capacitación y desarrollo de la persona que funja como administrador del sistema, por las labores de vigilancia y monitoreo constante, exigidas.

Finalmente, se concluye que, vincular las funciones de inteligencia a las de comunicación social, para que por medio de la minería textual se pueda develar de manera global, lo que sucede en un segmento de negocios, es la meta que persigue esta investigación.

6.2 Recomendaciones.

Con la idea de replicar esta propuesta a otros sectores industriales, se recomienda:

a. Para la aplicación del algoritmo, es necesario realizar las siguientes tareas:

1. La definición y delimitación de empresas por muestrear
2. El análisis detallado de un conjunto representativo de diferentes noticias, para determinar los elementos a considerar como esenciales
3. El acopio de los verbos y referentes de objetos más usuales en el ámbito bajo estudio
4. La definición de los términos de conexión (previos y extensores) de los elementos antes mencionados
5. La realización de pruebas manuales sobre un conjunto de noticias tipo para fines de comprobación
6. La mejora del sistema mediante la retroalimentación que pueda obtenerse de la implementación de las aplicaciones de cómputo

b. Para la explotación de los datos se requiere de las actividades siguientes:

1. Determinar las preguntas genéricas que se pretenden responder
2. Configurar los hechos que al publicarse evidencian las respuestas a las preguntas
3. Especificar las acciones principales que aparecen en los eventos que se desean detectar
4. Registrar los referentes de objetos que generalmente se presentan en los eventos mencionados previamente
5. Actualizar los criterios para responder a las preguntas que se pretenden plantear, para que la generación automática de los comandos SQL sea efectiva.

6.3 Trabajo futuro.

Entre las tareas por realizar se puede plantear que se hace necesario:

Para el algoritmo propuesto:

- Agregar nuevas locuciones, participios y expresiones indicativas de verbo, pues como cualquier ente vivo, el léxico nunca termina de crecer. Es notable la variedad de expresiones encontradas a lo largo del trabajo, por lo cual el mantenimiento a los catálogos correspondientes, será una tarea continua y continuada.
- Aplicar el software construido a partir del algoritmo a una muestra mayor de noticias del mismo sector, para determinar si la efectividad y la razón principal de falla son similares a los resultados encontrados.
- Determinar el grado de aplicabilidad a muestras de noticias empresariales de sectores distintos al educativo superior privado.
- Medir los resultados de aplicar el algoritmo a noticias publicadas en la prensa para determinar el grado de aplicabilidad de la solución.
- Buscar áreas de aplicación distintas donde los textos a procesar cuenten con estructura similar a las noticias.

Para el sistema de inteligencia competitiva:

- Validar su grado de utilidad con usuarios ejecutivos reales ante una muestra mayor de noticias procesadas.
- Desarrollar cuestionamientos adicionales a resolver por el sistema, pues las necesidades de los tomadores de decisiones se amplían y/o profundizan.
- Para una implementación de llave en mano, se requiere de la designación, capacitación y seguimiento del administrador del sistema, quien será el responsable de mantener actualizados los catálogos.

Finalmente, cualquier diseño de solución es susceptible de mejorar luego de enfrentarse a la realidad que supone la construcción de los elementos diseñados. Con base a lo anterior, se requerirá realizar actividades de ajuste en todos y cada uno de sus componentes.

6.4 Impacto social.

Actualmente, una de las principales aspiraciones de la empresa es convertirse en socialmente responsable. Las universidades particulares por su origen, convierten esa aspiración en una más de sus funciones. Para que el impacto sea mayor, cada entidad educativa ha de buscar las actividades que sus homólogas comunican que han realizado y los resultados logrados, para luego del ajuste particular que el ambiente requiere, reproducir las mejores prácticas que impulsen el desarrollo social. Para conseguirlo, un sistema como el propuesto permitirá a los directivos del área correspondiente: conocer, comparar y rediseñar su estrategia de participación y beneficio a las distintas comunidades que le atañen.

BIBLIOGRAFÍA Y REFERENCIAS.

- [1] MILANO, Antonio. Resolución de problemas y toma de decisiones. Editorial Macchi, 1993, Capitulo III, pag. 120-123.
- [2] CUBILLO, Julio. Techno-Economic Intelligence, Viena, UNIDO 1991
- [3] SOSA, Marisel y COBO, Pedro Humberto. Herramientas para la inteligencia empresarial. DISAIC. Disponible en: <http://www.nuevaempresa.cu/prog%20prel.htm>
- [4] PEÑA, Gustavo. La inteligencia tecnoeconómica: una ventana hacia el futuro. Espacios. Vol.13 (2) 1992. Disponible en: <http://www.revistaespacios.com/a92v13n02/20921302.html>
- [5] ARANO, Silvia y ASSTEGIANTE, Silvana. Repensando los procesos tecnicos desde la inteligencia organizacional en la universidad. CENTRO REGIONAL DE NUEVAS TECNOLOGIAS DE INFORMACION. CURSO: Inteligencia Organizacional, 26-28 de abril de 1999.
- [6] OROZCO, E. Preguntas y respuestas sobre la inteligencia empresarial. Disponible en: <http://www.nuevaempresa.cu/documentos/1preguntas.pdf>
- [7] GARCÍA, Susana. La inteligencia corporativa para el éxito empresarial. Cienc Inform 1998; 29(2):3-8.
- [8] HALAL, W. Organizational intelligence: What is it, and how can managers use it? Disponible en: <http://www.strategy-business.com/briefs/97413.html>
- [9] TOMIOKA, A. A corporate intelligence model: a new paradigm. Disponible en: [http://www.nacsis.ac.jp/sokuho/ARTICLES/NCID/an1004028x/19901100.3\(1\).html](http://www.nacsis.ac.jp/sokuho/ARTICLES/NCID/an1004028x/19901100.3(1).html)
- [10] TORRES, Ania. El profesional de la información en la inteligencia organizacional. Acimed Vol 10 05 2002. Disponible en: http://www.bvs.sld.cu/revistas/aci/vol10_5_02/aci02502.htm
- [11] DEAGOSTINI, Adriana y CORMENZANA, Fernando. Rol de las nuevas tecnologías de la información y la comunicación en la inteligencia organizacional aplicada a una pequeña y mediana empresa. Disponible en: <http://www.crnti.edu.uy/05trabajos/inteligencia/EnsayoIO.doc>
- [12] KAHANER, Larry. Competitive Intelligence. How to Gather, Analyze, and use Information to Move your Business to the top. Chapter 11: How the Japanese Perform Competitive Intelligence. 1997.
- [13] SCIP, Sociedad de Profesionales de Inteligencia Competitiva. SCIP 2005. <http://www.scip.org>
- [14] WALLE, A. "From Marketing Research to Competitive Intelligence: Useful Generalization or Loss of focus?." Management Decision. 1999

- [15] PUZZLE, Revista PUZZLE, DL. Número: B-45.448-03 ISSN 1696-8573, [http://www.revista-puzzle.com/Nota de prensa/20040510 PUZZLE - NOTA DE PRENSA - LA INTELIGENCIA COMPETITIVA EN LAS MULTINACIONALES CATALANAS.PDF](http://www.revista-puzzle.com/Nota%20de%20prensa/20040510%20PUZZLE%20-%20NOTA%20DE%20PRENSA%20-%20LA%20INTELIGENCIA%20COMPETITIVA%20EN%20LAS%20MULTINACIONALES%20CATALANAS.PDF)
- [16] GILAD, Benjamin y GILAD, Tamar. The Bussiness Intelligence System: A New Tool for Competitive Advntage. New York, AMACON. 1988
- [17] LAGERSTAM, Catharina. The Theory of Bussiness Intelligence. The Intelligence Process. Ponencia del Seminario "Information Industries Knowledge". Dubroknic, Yugoeslavia. 1989.
- [18] VILARAGUT L, Carro J. Internet como fuente para la inteligencia corporativa. Posibilidades y limitaciones. Disponible en: <http://www.nuevaempresa.cu/documentos/4internet.pdf>
- [19] ESTIVILL-CASTRO, Vladimir. Minería de Datos. LANIA. Año 6, Vols. 21 y 22, Otoño - Invierno 1997. Disponible en: <http://www.lania.mx/biblioteca/newsletters/1997-otono-invierno/mineria.html>
- [20] INFLEXA, Minería de datos. Disponible en: <http://www.inflexa.com/jsp/template.jsp?pag=mineria-datos.htm&mnu=mnu-mineria.htm>
- [21] VICENTE, Ma. Dolores. La aplicación de agentes inteligentes a la minería de web. Estudios de Información y Documentación de la UOC. Disponible en: <http://www.uoc.edu/web/esp/art/uoc/vicente0302/tfc/index1.html>
- [22] CHAKRABARTI 2000.
- [23] MARIN, María José, et al. DATA MINING EN RELACION A LA DOCUMENTACIÓN PERIODÍSTICA. Disponible en: <http://personales.upv.es/ccarrasc/doc/2003-004/DMPeriodistico/TREBALLSRP.htm>
- [24] MONTES-Y-GÓMEZ, Manuel. Minería de texto: Un nuevo reto computacional. Laboratorio de Lenguaje Natural, Centro de Investigación en Computación, Instituto Politécnico Nacional. Disponible en: <http://ccc.inaoep.mx/~mmontesg/publicaciones/2001/MineriaTexto-md01.pdf>.
- [25] TAN, Text Mining: The state of the art and challenges, Proc. of the Workshop Knowledge Discovery from advanced Databases PAKDDD-99, Abril 1999.
- [26] BORDONI, Luciana y D'AVANZO, Ernesto. Perspectivas para la integración de la minería de textos y la gestión del conocimiento. ENEA. Disponible en: <http://www.jrc.es/home/report/spanish/articles/vol68/ICT2S686.html>
- [27] STATSOFT. Text Mining. Disponible en <http://www.statsoft.com/textbook/sttextmin.html>
- [28] BOLASCO, Sergio. WG3: Text mining and knowledge management applications. Disponible en: <http://nemis.cti.gr/wg3.htm>
- [29] Technology Policy and Assessment Center (TPAC) at Georgia Institute of Technology. TEXT MINING. Disponible en: [http://www.asdl.gatech.edu/research_teams/pdf/2002/Text Mining 20Sum.doc](http://www.asdl.gatech.edu/research_teams/pdf/2002/Text%20Mining%20Sum.doc)

- [30] RAVINDRAN, Balaji. TEXT MINING APPLICATIONS TO DATA MINING.
Disponible en: <http://www.highbrew.com/whitepapers/text-mining-in-marketing.pdf>
- [31] Second Moment. Evaluating Text Mining Applications
Disponible en:
<http://www.secondmoment.org/Stats in Industry/Evaluating Text Mining Applications.htm>
- [32] TÉLLEZ-VALERO y otros. Clasificación automática de textos de desastres naturales en México. Disponible en: Documento1.pdf
- [33] GARCÍA-MENIER, E., Un Sistema para la Clasificación de Notas Periodísticas, Proc. of the Symposium Internacional de Computación CIC-98, México, D. F., 1998.
- [34] MÉNDEZ-TORREBLANCA, A., MONTES-Y-GÓMEZ M. y LÓPEZ-LÓPEZ, A. A Trend Discovery System for Dynamic Web Content Mining. Proc. of the XI International Conference on Computing, Mexico City, México, November 2002.
- [35] GUZMÁN, A. Hallando los temas principales en un artículo en español. (1997) Soluciones Avanzadas. Vol. 5, núm. 45, pág. 58ff, I parte, 15 de Julio; II parte vol. 5, núm. 49, pág. 66ff, 15 de septiembre.
- [36] Manuel Montes y Gómez, Aurelio López López. Nominalizations in titles: a way to extract document details. Proc. CIC-98, Symposium Internacional de Computación, November 11-13, 1998, Mexico D.F., pp. 396-404.
- [37] MONTES Y GÓMEZ, Manuel, GELBUKH, Alexander y LÓPEZ LÓPEZ, Aurelio. Mining the news: trends, associations, and deviations. Computación y Sistemas, Revista Iberoamericana de Computación, ISSN 1405-5546 (included in the Index of CONACyT), Vol. 5 N 1, 2001, pp. 14-24
- [38] MONCADA, M. A. (2001). Fines y medios de Comunicación Social.
<http://teologia.ucsc.cl/noticias/300805/mmoncada.pdf>
- [39] PIÑUEL, J. L. (1999). Comunicación y Terrorismo. Universidad Complutense de Madrid.
- [40] RODRÍGUEZ, Juan Carlos, Dr. Centeno (2005-2006) Aproximación Teórica a la Publicidad. El Sistema Publicitario, Integrantes e Interconexiones; Razón y Palabra, N° 48, Facultad de Comunicación, Universidad de Sevilla, España. Descargado de <http://www.razonypalabra.org.mx/actual/jrodriguez.html#au#au>
- [41] RODRIGO, Miguel. La construcción de la noticia, Paidós, Comunicación, Barcelona, 1996, p.113.
- [42] DE FONTCUBERTA, Mar. "Pauta y calidad informativa", Cuadernos de Información N°13, Escuela de Periodismo, Pontificia Universidad Católica de Chile, Santiago, 1999, pp. 61-69.
- [43] ROZAS, Eliana. "La selección noticiosa, entre la importancia y el interés", Cuadernos de Información N°12, Escuela de Periodismo, Pontificia Universidad Católica de Chile, Santiago, 1997, pp. 20-25.

- [44] ROMAÑACH, J. Utilización de Internet por parte de las personas con discapacidad con fines de comunicación social. Un informe empírico. En: Delgado, B. y Fernández, B. (coord.): Nuevos Conceptos de Comunicación. Madrid: Universidad Complutense. 1999.
- [45] SHABB, Verónica. Formatos periodísticos para prensa escrita. 2004.
http://www.icuar.com.ar/documentos/todos/Generos_periodisticos_color.pdf
- [46] MARTINEZ J.L. MARTINEZ Albertos, Guiones de clase de Redacción Periodística, Pamplona, 1962.
- [47] CASTRESANA, Christian. ¿Cómo redactar una buena nota de prensa? 2005.
Descargado de: <http://www.noticias.info/aprenda/portada.shtm>
- [48] MORIELLO, Sergio. ISSN: 1579-0223, Cyborgs: hijos de la mente; Red Científica, <http://www.redcientifica.com/doc/doc199908080011.html>
- [49] DEL VALLE, Felix. Análisis y tratamiento documental en medios de comunicación impresos en la era digital. Universidad Complutense de Madrid.
<http://www.ucm.es/info/multidoc/multidoc/cursos/verano/material/FELIX DEL VALLE/fvalleEV4.htm>
- [50] COLLE, Raymond. "Análisis lógico de hechos noticiosos", Revista latina de comunicación social, Num. 27, Marzo 2000, La Laguna (Tenerife). 2000 URL: <http://www.ull.es/publicaciones/latina/aa2000tma/126colle.html>
- [51] M. Montes- y-Gómez 1 , A. Gelbukh 2 , and A. López- López 1 Text Mining at Detail Level using Conceptual Graphs
- [52] Boytcheva, Dobrev y Angelova (2001), CGExtract: Towards Extraction of Conceptual Graphs from Controlled English. Lecture Notes in Computer Science 2120, Springer 2001.
- [53] Rassinoux A-M, Baud RH, Scherrer J-R. A multilingual analyser of medical texts. In: Second International Conference on Conceptual Structures (ICCS 94), University of Maryland, 1994.
- [54] J. Hess and W. Cyre (1999), "A CG-based Behavior Extraction System," Proc. Seventh International Conference on Conceptual Structures, Blacksburg, VA, 127-139, July 12-15, 1999. 1999 Springer-Verlag
- [55] Petersen Ulrik. Creation in Graphs. Extracting conceptual structures from Old Testament texts. Aalborg University, Denmark. May 18-22, 2003
- [56] J.F. Sowa: Conceptual Structures: Information Processing in Mind and Machine. Reading, MA: Addison-Wesley Publishing Company 1984.

- [57] J.F. Sowa: *Conceptual Structures: Information Processing in Mind and Machine*. Reading, MA: Addison-Wesley Publishing Company 1984.
- [58] J. Hess and W. Cyre (1999), "A CG-based Behavior Extraction System," Proc. Seventh International Conference on Conceptual Structures, Blacksburg, VA, 127-139, July 12-15, 1999. 1999 Springer-Verlag
- [59] N. Sebastián, M.A. Martí, M.F. Carreiras y F. Cuetos (2000) *LEXESP: Léxico Informatizado del Español*, Edicions de la Universitat de Barcelona.
- [60] L. Ruiz. *Relaciones categoriales de las locuciones adverbials*. Universitat de València.
- [61] Esquivel, Ismael (2006). *Tratamiento automático de noticias empresariales en la Web*. 4º. Congreso Internacional de la CENIDET-IEEE.

A n e x o s

Clave	Descripción	Página
A	Empresas	
B	Objetos	
C	Verbos	
D	Términos	
E	Mapeo de elementos básicos	
F	Locuciones no verbales	
G	Locuciones verbales	
H	Lista de participios de presente	
I	Diccionario de datos	
J	Resultados del procesamiento de noticias	
K	Resultados de la sustitución de elementos en noticias	
L	Reglas para elección del verbo principal	

Anexo A. Empresas

Empresa	Liga
UDLA	http://www.udlap.mx/noticias/SalaPrensa.aspx
UCC	http://www.ver.ucc.mx/tablero/
IBERO	http://www.iberro.edu.mx/actividades/comunicados/default.html
IBERO	http://www.iberro.edu.mx/actividades/salaprensa/agenda/default.html#P
ULSA	http://www.ulsu.edu.mx/noticias/listadoNoticias.php
EBC	http://www.ebc.mx/contenido/acercade/ACNO01.htm
UIC	http://www.uic.edu.mx/uic/noticias.aspx
UNITEC	http://www.unitec.mx/portal/page?_pageid=537,905609&_dad=portal&_schema=PORTAL
UAG	http://www.uag.mx/202/noticias_aldia.htm
UPAEP	http://web.upaep.mx/PrensaCom/boletin.asp
UPAEP	http://web.upaep.mx/PrensaCom/boletinDeportivo.asp
ITESM-PUE	http://www.pue.itesm.mx/nueva/principal.html
CETYS-MXL	http://www.mxl.cetys.mx/noticias.php
CUGS	http://www.cugs.edu.mx
ITESO	http://www.iteso.mx/publico/comunicacion/index.jsp
ANAHUAC	http://www.anahuac.edu.mx/
LASALLE	http://intranet.delasalle.edu.mx/nueva/intranet3/noticias/listado.php
UAM	http://www.uam.mx/avisos/index.html
UDEM	http://www.udem.edu.mx/agencia/udem/
UP	http://www.up.mx/
UR	http://expresion.ur.mx/
UVM	http://www.uvmnet.edu/

ANEXO B. Términos asociados a referentes de los objetos

acceso	colección	estudios	muestra	recorrido
acervo	coloquio	etapa	mundo	regional
acreditación	comida	evento	noche	rendición
actividad	competencia	eventos	nombramiento	respuesta
actuación	concierto	examen	obra	reunión
acuerdo	conclusión	exhibición	observación	semana
agrupación	concurso	exposición	obtención	seminario
alianza	conferencia	fase	oferta	Servicio
altar	conformación	feria	opción	símbolo
análisis	congreso	festival	operación	simposio
apertura	constancia	fiesta	organización	simposium
aportación	convenio	final	panel	situación
apoyo	convenios	firma	película	taller
asamblea	convivencia	formación	peso	temas
aula	convocatoria	foro	placa	temporada
beca	creación	graduación	planeación	titulación
brindis	credencial	homenaje	planta	toma_de_protesta
cadena	crítica	ideal	plataforma	trabajo
campamento	cuarto	impartición	plática	traducción
campeón	curso	importancia	potencia	victoria
campeonato	deporte	inauguración	premiación	videoconferencia
capacitación	desayuno	informe	premio	vinculación
carta	despedida	iniciación	premios	virus
casa	día	instalaciones	presentación	vocación
cátedra	dictamen	institución	producción	
celebración	diplomado	jóvenes	programa	
cena	disposición	juramento	propiedad	
centro	doctorado	la conferencia	propuesta	
ceremonia	documental	libro	proyecto	
certificación	donación	licenciatura	racha	
charla	el concierto	maestría	recepción	
ciclo	encuentro	medalla	recibimiento	
coctel	escuela	mención	reconocimiento	

Anexo C. Relación de Verbos

Infinitivo	Raíz	Tipo	Infinitivo	Raíz	Tipo
aceptar	accept		hablar	habl	
afirmar	afirm		impartir	impart	
ampliar	ampl		inaugurar	inaugur	
analizar	anali		informar	inform	
anunciar	anunci		iniciar	inici	
apabullar	apabull		instar	inst	
apoyar	apoy		invertir	inv	
avaluar	aval		invitar	invit	
ayudar	ayud		ir	*	
brindar	brind		llevar	llev	
buscar	bus		lograr	logra	
celebrar	celebr		obtener	obt	
clausurar	clausur		opinar	opin	
comenzar	comenz		organizar	organiz	
concebir	conc		otorgar	otorg	
concluir	conclu		permanecer	permane	
consolidar	consolid		planificar	plan	
contar	*		posicionar	posicion	
convenir	conv		premiar	premi	
convivir	convivi		presentar	present	
crear	cre		pretender	pretend	
dedicar	dedi		producir	produ	
despedir	desp		proyectar	proyect	
develar	devel		realizar	reali	
donar	don		recibir	recib	
efectuar	efectu		reconocer	reconoc	
enfrentar	enfrent		rendir	*	
engalanar	engalan		retornar	retorn	
entregar	entreg		reunir	reun	
establecer	establec		seleccionar	seleccion	
estudiar	estudi		ser	*	
exponer	exp		trabajar	trabaj	
expresar	expres		tramitar	tramit	
firmar	firm		unir	uni	
ganar	gan		venir	*	

Infinitivo	raíz	tipo	infinitivo	raíz
acceder	acced		fijar	fij
aceptar	accept		firmar	firm
acordar	acord		fortalecer	fortalec
Acreditar	acredit		ganar	gan
acudir,	acud		haber	hab
afirmar	afirm		hablar	habl
albergar	alberg		impartir	impart
ampliar	ampl		implementar	implement
analizar	anali		inaugurar	inaugur
anunciar	anunci		informar	inform
apabullar	apabull		iniciar	inici
apoyar	apoy		insistir	insist
asegurar	asegur		instar	inst
asistir	asist		invertir	invert
atacar	ata		invertir	inv
avalar	aval		invitar	invit
ayudar	ayud		ir	*
bailar	bail		llevar	llev
brindar	brind		lograr	logr
buscar	bus		mencionar	mencion
calificar	calific		montar	mont
celebrar	celebr		necesitar	necesit
clasificar	clasifi		obedecer	obedec
clausurar	clausur		obtener	obt
comenzar	comenz		oficiar	ofici
comprar	compr		ofrecer	ofrec
concebir	conc	Auxiliar	opinar	opin
concluir	conclu		organizar	organiz
conocer	cono		otorgar	otorg
conseguir	cons	Auxiliar	permanecer	permane
consolidar	consolid	Auxiliar	planificar	plan
contar	*		platicar	plati
contemplar	contempl		poder	pod
continuar	continú		poner	pon
contratar-	contrat		posicionar	posicion
convenir	conv	Auxiliar	premiar	premi
convivir	convivi		preparar	prepar
crear	cre		presentar	present
dar	d		pretender	pretend
decir	dec		producir	produ
decorar	decor		proteger	proteg

Infinitivo	raíz	tipo	Infinitivo	raíz	tipo
dedicar	dedi		proyectar	proyect	
dejar	dej		realizar	reali	
deleitar	deleit		recibir	recib	
desarrollar	desarroll		reconocer	reconoc	
despedir	desp		rendir	*	
destinar	destin		retornar	retorn	
develar	devel		reunir	reun	
discutir	discut		seleccionar	seleccion	
donar	don		señalar	señal	
efectuar	efectu		ser	*	
elaborar	elabor		solucionar	solucion	
encomendar	encomend		temer	tem	
enfocar-	enfoc		tener	ten	
enfrentar	enfrent		trabajar	trabaj	
engalanar	engalan		tramitar	tramit	
entregar	entreg		transmitir	transmit	
establecer	establec		tributar	tribut	
estar	est		unir	uni	
estudiar	estudi		venir	*	
exponer	exp		versar	vers	
expresar	expres		visitar	visit	

Anexo D. Elementos de conexión y delimitación.

PREPOSICIONES

a	alrededor de	en nombre de	según
a base de	amén de	en pos de	sin
a cambio de	ante	en relación a	so
a causa de	aparte	en torno a	sobre
a comienzos de	aparte de	en vez de	tras
a costa de	bajo	en vías de	a costa del
a diferencia de	con	entre	a mitad del
a eso de	con objeto de	excepto	a partir del
a excepción de	con respecto a	frente a	a pesar del
a favor de	con tal de	gracias a	a través del
a n de	contra	hacia	al
a fuerza de	de	hasta	alrededor del
a largo de	de acuerdo con	junto a	amén del
a lo largo de	debido a	junto con	con respecto al
a modo de	dentro	lejos de	de cara al
a nombre de	dentro de	mediante	debajo del
a partir de	desde	pa	debido al
a pesar de	después de	para	del
a punto de	durante	pese a	dentro del
a raíz de	en	por	en cuanto al
a través de	en ausencia de	por debajo de	en medio del
acerca de	en calidad de	por delante de	en torno al
además de	en contra de	por encima de	en virtud del
al cabo de	en cuanto a	por medio de	frente al
al contrario de	en favor de	por obra y gracia de	junto al
al frente de	en lugar de	por parte de	pese al
al lado de	en medio de	respecto a	por encima del
al margen del	en menos de	respecto de	respecto al
al menos	en mitad de	salvo	respecto del

CONJUNCIONES

a saber	a medida que	de ahí que	para que
e	a pesar de que	de aquí que	pese a que
es decir	adonde	de modo que	por más que
esto es	ahora que	de tal manera que	porque
mientras	al igual que	debido a que	pues
ni	al mismo tiempo que	desde que	puesto que
no obstante	al tiempo que	después de	que
o	antes de que	donde	según que
o bien	así como	en caso de que	si
o sea	así es que	en cuanto	si bien
pero	así que	en el caso de que	siempre que
pese a que	aun	en tanto que	sino que
sin embargo	aunque	entonces	tal como
sino	bien que	hasta que	tal y como
sino que	cada vez que	mas	tan pronto como
u	como	mientras	tanto como
y	como quiera que	mientras que	una vez que
a la par que	cuando	mientras tanto	ya que
a la vez que	dado que	ni que	

ARTICULOS

el
la
lo
los
las

ADVERBIOS

a borbotones	convenientemente	estrechamente	poco a poco
a bordo	cortésmente	estrictamente	poderosamente
a buen seguro	cotidianamente	estúpida	políticamente
a cambio	crudelísimamente	evidentemente	por aquel entonces
a ciegas	cruelmente	evolutivamente	por casualidad
a continuación	crónicamente	exactamente	por cierto
a cubierto	cuando menos	excepto	por completo
a cuestras	cuanto	excesivamente	por consiguiente
a diario	cuanto antes	exclusivamente	por ejemplo
a fin de cuentas	cuidadosamente	explícitamente	por el contrario
a flote	culturalmente	extraordinariamente	por encima
a fondo	cum grano salis	extremadamente	por entero
a gusto	curiosamente	famosamente	por favor
a la larga	cáusticamente	fatalmente	por fin
a la perfección	cómodamente	favorablemente	por fortuna
a la postre	de churro	felizmente	por igual
a la vez	de acuerdo	fijamente	por lo común
a las mil maravillas	de antemano	finalmente	por lo menos
a lo lejos	de bruces	fotográficamente	por lo tanto
a lo mejor	de buen grado	fraccionalmente	por menos
a lo que parece	de cara	francamente	por otra parte
a lo sumo	de costumbre	frente	por otro lado
a medias	de cualquier forma	fuera	por si acaso
a menudo	de cuando en cuando	fuertemente	por sorpresa
a muerte	de golpe	fundamentalmente	por supuesto
a ojos vistas	de hecho	futbolísticamente	por tanto
a partir de ahí	de inmediato	fácilmente	por todo lo alto
a plazos	de largo	física	por último
a plomo	de lo contrario	generalmente	posteriormente
a priori	de maravilla	genialmente	precipitadamente
a propósito	de menos	geográficamente	precisamente
a puntapiés	de momento	gozosamente	preferiblemente
a punto	de muerte	gratis	prematuramente
a ratos	de nuevo	habitualmente	previamente
a sabiendas	de otro lado	harto	primero
a salvo	de oídas	hasta	probablemente
a sangre y fuego	de par en par	herméticamente	profesionalmente
a secas	de paso	hipócritamente	profundamente
a solas	de pleno	honestamente	profusamente
a tiempo	de por vida	hoy	progresivamente
a tiro limpio	de pronto	hoy día	pronto
a toda prisa	de repente	hoy por hoy	propiamente
a todas horas	de rodillas	humildemente	proporcionalmente
a todo pasto	de siempre	igual	prudentemente
a veces	de súbito	igualmente	prácticamente
abajo	de tanto en tanto	ilegalmente	próximamente
abreviadamente	de tarde en tarde	impepinablemente	publicitariamente
absolutamente	de todas formas	impunemente	pues bien
acaso	de todos modos	in extremis	puntualmente
acertadamente	de veras	incesantemente	puramente
actualmente	de verdad	incluso	púdicamente
adelante	de vez en cuando	incomparablemente	quizá

además	debajo	incomprensiblemente	quizás
adentro	debidamente	inconscientemente	rara vez
adicionalmente	debido	indeciblemente	raramente
admirablemente	decididamente	individualmente	razonablemente
afortunadamente	decorosa	industrialmente	realmente
ahora	definitivamente	inefablemente	recien
ahora bien	del todo	inesperadamente	recientemente
ahora mismo	delante	inevitablemente	recién
ahí	demasiado	inexorablemente	recíproca
al alcance de la mano	democráticamente	infinitamente	reiteradamente
al cabo	densamente	inmediatamente	relativamente
al cien por cien	dentro	inocentemente	repentinamente
al contado	deprisa	insospechada	respectivamente
al contrario	desde entonces	insólitamente	ridículamente
al fin	desde luego	intelectualmente	rigurosamente
al fin y al cabo	desgraciadamente	inteligentemente	rápidamente
al final	desordenadamente	internacionalmente	salvo
al instante	despacio	intrínsecamente	samzin
al menos	después	irremediablemente	secretamente
al momento	detalladamente	irremisiblemente	secularmente
al máximo	detrás	jamás	seguramente
al parecer	diagonalmente	justamente	seguro
al pie de la letra	difícilmente	justo	semimilagrosamente
al poco	dignamente	lamentablemente	sencillamente
al principio	directamente	largo	seriamente
al revés	discretamente	legalmente	sexualmente
alegremente	discretísimamente	lejos	si acaso
algo	distintivamente	lentamente	si bien se mira
allá	doblemente	levemente	sic
allí	dolorosamente	libremente	siempre
alrededor	drásticamente	ligeramente	simplemente
altamente	dudosamente	literalmente	simultáneamente
alternativamente	dulcemente	literariamente	sin ambages
amargamente	efectivamente	lo menos	sin duda
amistosamente	eficazmente	lo mismo	sin embargo
ampliamente	electroquímicamente	luego	siquiera
anoche	electrónicamente	lógicamente	soberanamente
anteayer	eléctricamente	magníficamente	sobre ruedas
anteriormente	en absoluto	mal	sobre todo
antes	en alto	mano a mano	sobremanera
antes al contrario	en armas	mañana	solamente
antiguamente	en boga	mecánicamente	solo
antiguo	en breve	medio	Sonrientemente
aparentemente	en broma	mejor	Sorprendentemente
aparte	en cambio	menos	Subjetivamente
apenas	en cierto modo	menos mal	Sucesivamente
aproximadamente	en concreto	menos mal que	Suficientemente
aquí	en conjunto	mensualmente	Sumamente
arriba	en consecuencia	meramente	Sumariamente
artificialmente	en contra	metafísicamente	Supuestamente
asimismo	en cuanto	mientras	Sustancialmente
astutamente	en cuerpo y alma	mientras tanto	Sí
así	en definitiva	milagrosamente	Sólo
así como	en diagonal	milimétricamente	Tal

así pues	en efecto	mismo	tal cual
así y todo	en el acto	modestamente	tal vez
atrás	en exceso	moralmente	También
aun	en fin	mortalmente	Tampoco
averías	en general	mucho	Tan
ayer	en gran medida	mutuamente	tan solo
aún	en medio	muy	Tanto
bastante	en nada	más	Tarde
bien	en parte	más aún	Temporalmente
biológicamente	en particular	más bien	Temprano
blanco	en principio	más o menos	termodinámicamente
brevísimamente	en privado	más y más	Terriblemente
bruscamente	en profundidad	nada	Textualmente
bueno	en realidad	naturalmente	Teóricamente
básicamente	en rigor	necesariamente	Tibiamente
calladamente	en seguida	nerviosamente	Todavía
casi	en serio	ni mucho menos	todo lo contrario
cautelosamente	en su defecto	ni siquiera	Torpemente
cerca	en tanto	ni tantico así	Total
cerca de	en todo caso	no	Totalitariamente
	en un abrir y cerrar de		
ciertamente	ojos	no obstante	Totalmente
cierto	en vano	normalmente	Tradicionalmente
claramente	en verdad	notablemente	Tranquilamente
clarito	en vilo	nuevamente	Tremendamente
claro	en zigzag	nunca	Térmicamente
clásicamente	encima	ocultamente	Típicamente
cobarde	enfrente	oficialmente	un poco
codo a codo	enigüei	orgullosamente	Universalmente
colectivamente	enormemente	ostensiblemente	uno a uno
como consecuencia	enseguida	otrosí	Usualmente
como dios manda	enteramente	para colmo	uña y carne
como es debido	entonces	paradójicamente	Vagamente
como mínimo	entre tanto	paralelamente	Veladamente
Como si tal cosa	entretanto	parcialmente	Verbigracia
comparativamente	es decir	particularmente	Verdaderamente
completamente	escrupulosamente	pasablemente	Veri
comúnmente	esencialmente	paulatinamente	Visto
con fervor	espacialmente	peor	Voluntariamente
concretamente	especialmente	perfectamente	Vulgarmente
condicionalmente	específicamente	permanentemente	y todo
confiadamente	espléndidamente	personalmente	Ya
constantemente	espontáneamente	plácidamente	Últimamente
contradictoriamente	estratégicamente	poco	Únicamente

DEMOSTRATIVOS

tales	semejantes	
Tal	semejante	
estas	Esas	aquellas
Esta	Esa	aquella
estos	Esos	aquellos
este	Ese	aquel

INDEFINIDOS

alguno	Demás	ninguno	uno
bastante	Diferente	otro	varios
cada	Distinto	poco	unas
cierto	Escaso	propio	una
cualquiera	Escasísimo	sendos	unos
cuanto	Mismo	tanto	un
demasiado	Mucho	todo	

INTERROGATIVOS

qué
cuanta
cuántas
cuánto
cuántos

POSESIVOS

mis	nuestra	tuyas	vuestro
mi	nuestros	tuya	suyas
míos	nuestro	vuestras	suya
mío	tus	vuestra	suyos
nuestras	tu	vuestros	suyo

PRONOMBRES

conmigo	ellos	me	sí
consigo	la	mí	te
contigo	las	nos	ti
él	le	nosotras	tú
ella	les	nosotros	vosotras
ellas	lo	os	vosotros
ello	los	se	yo

ANEXO E. Mapeo de elementos básicos.

A continuación se presentan los términos del español que generalmente se asocian a cada uno de los elementos básicos. Para abreviar, se indican términos como:

[NP] = Nombre propio

[NC] = Nombre común

Previos = Términos que anteceden al nombre del elemento

Extensores = Términos que extienden al nombre del elemento

Delimitadores = Símbolos ortográficos que indican el fin del elemento

AGENTE

previos	[NP]	extensores	[NP]
	[NC]		
al		a cargo de	
como		a través de	
de acuerdo a		a través del	
de acuerdo con		como	
nuestra(o)		con	
por		conformado por	
organizado(a) por		de	
el		del	
la		e	
los		en	
las		nuestra(o)	
		para	
		perteneciente a	
		que dirige	
		y	

OBJETO

previos	[NC]	extensores	[NC] [NP]	Delimit	previos	OBJETO DET.	Delimit
a		de		[coma]	a cargo de		[coma]
a esta		del		[punto]	a nivel		[punto]
a este		e			al		
al		y			cada		
como					como		
con su					con		
del					de		
en					del		
es uno de					denominado		
esta					desde		
su					durante		
sus					el pasado		
un					el próximo		
una					en		
					en el tema de		
					este		
					integrado por		
					para		
					por		
					que		
					sobre		
					titulado		
					y		

PACIENTE

previos	[NP] [NC]	extensores	[NP]
a		a cargo de	
al		a través de	
apoya a		con	
con		conformado por	
consiguió de		de	
contarán con		del	
del		perteneciente a	
en en apoyo a		y	
frente a			
fue reconocido con			
obtuvo de			
para			
para beneficio de			
por el			
recibió de			

EFEECTO

previos	términos	delimitador
que permit		[coma]
que facilit		[punto]
que condu		
que contribuy		
con el fin de		
con el propósito de		
con la idea de		
con el objetivo de		
con la finalidad de para		

CONTEXTO

previos	términos	delimitador
como parte de		[coma]
en el marco de		[punto]
en el contexto de		

INSTRUMENTO

previos	términos	delimitador
por medio de		[coma]
por		[punto]
auspiciado por		
patrocinado por		

ANEXO F. LOCUCIONES NO VERBALES

LOCUCIONES PREPOSICIONALES

a_falta_de	a_raíz_de	con_tal_de	en_menos_de	pese_al
a_base_de	a_través_de	de_acuerdo_con	en_mitad_de	por_debajo_de
a_cambio_de	acerca_de	de_cara_al	en_nombre_de	por_delante_de
a_causa_de	además_de	debajo_de	en_pos_de	por_encima_de
a_comienzos_de	al_cabo_de	debajo_del	en_relación_a	por_encima_del
a_costa_del	al_contrario_de	debido_a	en_torno_a	por_medio_de
a_diferencia_de	al_frente_de	debido_al	en_torno_al	por_obra_y_gracia_de
a_eso_de	al_lado_de	dentro_de	en_vez_de	por_parte_de
a_excepción_de	al_margen_del	dentro_del	en_vías_de	respecto_a
a_favor_de	al_menos	después_de	en_virtud_del	respecto_al
a_fin_de	alrededor_de	detrás_de	encima_de	respecto_de
a_fuerza_de	amén_de	en_ausencia_de	enfrente_de	respecto_del
a_largo_de	amén_del	en_calidad_de	frente_a	rumbo_a
a_lo_largo_de	antes_de	en_contra_de	frente_al	
a_mitad_del	aparte_de	en_cuanto_a	gracias_a	
a_modos_de	camino_de	en_cuanto_al	junto_a	
a_nombre_de	con_arreglo_a	en_favor_de	junto_al	
a_partir_de	con_objeto_de	en_lugar_de	junto_con	
a_pesar_de	con_respecto_a	en_medio_de	lejos_de	
a_punto_de	con_respecto_al	en_medio_del	pese_a	

LOCUCIONES CONJUNTIVAS

a_fin_de_que_si	bien_que	en_tanto_que	por_más_que
a_la_par_que	cada_vez_que	es_decir	porque
a_la_vez_que	como	esto_es	pues
a_medida_que	como_quiera_que	hasta_que	puesto_que
a_pesar_de_que	con_tal_que	igual...que	que
a_que	cuando	más...que	según_que
a_saber	dado_que	menos... que...que	si
ahora_que	de_ahí_que	mientras	si_bien
al_igual_que	de_aquí_que	mientras_que	siempre_que
al_mismo_tiempo_que	de_modo_que	mientras_tanto	sin_embargo
al_tiempo_que	de_tal_manera_que	ni_que	sino_que
antes_de_que	debido_a_que	no_obstante	tal_como
apenas	desde_que	o_bien	tal_y_como
así_como	después_de	o_sea	tan_pronto_como
así_es_que	en_caso_de_que	para	tanto_como
así_que	en_cuanto	para_que	una_vez_que
aunque	en_el_caso_de_que	pese_a_que	ya_que

LOCUCIONES ADVERBIALES

a_borbotones	al_contrario	de_repente	en_vano
a_bordo	al_fin	de_rodillas	en_verdad
a_buen_seguro	al_fin_y_al_cabo	de_siempre	en_vilo
a_cambio	al_final	de_súbito	en_zigzag
a_ciegas	al_instante	de_tanto_en_tanto	entre_tanto
a_continuación	al_menos	de_tarde_en_tarde	es_decir
a_cubierto	al_momento	de_todas_formas	hoy_día
a_cuestas	al_máximo	de_todos_modos	hoy_por_hoy
a_diario	al_parecer	de_veras	in_extremis
a_n_de_cuentas	al_pie_de_la_letra	de_verdad	lo_menos
a_ote	al_poco	de_vez_en_cuando	lo_mismo
a_fondo	al_principio	del_todo	mano_a_mano
a_gusto	al_revés	desde_entonces	menos_mal_que
a_la_larga	antes_al_contrario	desde_luego	más_aún
a_la_perfección	así_como	en_absoluto	más_bien
a_la_postre	así_pues	en_alto	más_o_menos
a_la_vez	así_y_todo	en_armas	más_y_más
a_las_mil_maravillas	cerca_de	en_boga	ni_mucho_menos
a_lo_lejos	codo_a_codo	en_breve	ni_siquiera
a_lo_mejor	como_consecuencia	en_broma	ni_tantico_así
a_lo_que_parece	como_dios_manda	en_cambio	
a_lo_sumo	como_es_debido	en_cierto_modo	para_colmo
a_medias	como_mínimo	en_concreto	poco_a_poco
a_menudo	como_si_tal_cosa	en_conjunto	por_aquel_entonces
a_muerte	con_fervor	en_consecuencia	por_casualidad
a_ojos_vistas	cuando_menos	en_contra	por_cierto
a_partir_de_ahí	cuanto_antes	en_cuanto	por_completo
a_plazos	cum_grano_salis	en_cuerpo_y_alma	por_consiguiente
a_plomo	de_"churro"	en_definitiva	por_ejemplo
a_priori	de_acuerdo	en_diagonal	por_el_contrario
a_propósito	de_antemano	en_efecto	por_encima
a_puntapiés	de_bruces	en_exceso	por_entero
a_punto	de_buen_grado	en_fin	por_favor
a_ratós	de_cara	en_general	por_fin
a_sabiendas	de_costumbre	en_gran_medida	por_fortuna
a_salvo	de_cualquier_forma	en_medio	por_igual
a_sangre_y_fuego	de_cuando_en_cua ndo	en_nada	por_lo_común
a_secas	de_golpe	en_parte	por_lo_menos
a_solas	de_hecho	en_particular	por_lo_tanto
a_tiempo	de_inmediato	en_principio	por_menos
a_tiro_limpio	de_largo	en_privado	por_otra_parte
a_toda_prisa	de_lo_contrario	en_profundidad	por_otro_lado
a_todas_horas	de_maravilla	en_realidad	por_si_acaso
a_todo_pasto	de_menos	en_rigor	por_sorpresa
a_veces	de_momento	en_seguida	por_supuesto
ahora_bien	de_muerte	en_serio	
ahora_mismisto	de_nuevo	en_su_defecto	
al_alcance_de_la_ma no	de_otro_lado	en_tanto	
al_cabo	de_oídas	en_todo_caso	
al_cien_por_cien	de_par_en_par	en_un_abrir_y_cerr ar_de_ojos	

Anexo G. Locuciones Verbales.

Locución verbal	Verbo
dar a conocer	informar
llevar a cabo	realizar
tomar en cuenta	considerar
dar inicio	iniciar
estar presente	convivir
acudir a realizar	realizar
de acuerdo a	acordar
dar forma	configurar
quedar sin	abstener

Casos especiales

Expresión	Verbo
lograr (Verbo)	(Verbo)
poder (Verbo)	(Verbo)
quedar (Verbo)	(Verbo)
invitar a (Verbo)	(Verbo)

Anexo H. Adjetivos y sustantivos con la terminación del participio de presente -ante, -ente, -iente

A

abastante	aislante	apasionante
aberrante	albicante	apelante
abracadabrante	algente	apellidante
absorbente	aliciente	aplaciente
abstergente	alienante	aplastante
abundante	almirante	apostante
accedente	alterante	aprehendiente
accidente	altisonante	aprendiente
acescente	altitonante	aprobante
acezante	altoparlante	aprovechante
acompañante	alucinante	aquiescente
acordante	alumbrante	arborescente
acrescente	amante	arbotante
actuante	ambiente	arcaizante
acudiente	ambivalente	ardiente
adoleciente	ambulante	armisonante
adolescente	anabolizante	arrastrante
adoptante	anarquizante	arrogante
adquiriente	andante	asaltante
adragante	anglicanizante	ascendente
adsorbente	anglohablante	ascendiente
adstringente	angloparlante	asfixiante
adyacente	animante	asintiente
adyuvante	antecedente	asonante
agarrante	anticipante	aspirante
agente	anticongelante	astringente
agitante	antideslizante	atediante
aglomerante	antidetonante	atemperante
aglutinante	antiespumante	atenuante
agonizante	antioxidante	atingente

agravante	anunciante	atorrante
aguante	apabullante	autosuficiente
aguardiente	apante	ayudante

B

bacante	beligerante	bienplaciente
bailante	benevolente	bienquiriente
bajante	bergante	bivalente
balante	besante	blasfemante
barbiponiente	bezante	boyante
barbipungente	bienandante	brabante
barbipuniente	bienhaciente	bramante
bastante	bienmereciente	brillante
batiente	bienoliente	brotante
becante	biempensante	buenameresciente

C

cabadelante	cociente	contaminante
cabestrante	coeficiente	contante
cabildante	coincidente	contendiente
cabrestante	colgante	contingente
cadente	colindante	contraalmirante
caladelante	colitigante	contracorriente
callante	colorante	contradicente
calmante	comandante	contradurmiente
cambiante	combatiente	contrafirmante
caminante	comediante	contraindicante
campante	comenzante	contralmirante
candente	comerciante	contrapariante
caneciente	comodante	contraproducente
cantante	compaciente	contrapuntante
cante	compareciente	contrastante
capelardente	complaciente	contratante

capitulante	comprante	contrincante
caradelante	comprobante	contundente
carburante	comulgante	convaleciente
cargante	comunicante	conveniente
caridoliente	concertante	convergente
casateniente	concupiscente	convincente
castellanohablante	conkursante	conviviente
caulescente	condecete	convulsionante
causahabiente	condescendiente	copiante
causante	condicionante	coribante
ceceante	condonante	corregente
cecuciente	conferenciante	correinante
celante	confesante	correspondiente
celebrante	conficiente	corriente
cementante	confidente	corroborante
cesante	confinante	corruscante
chispeante	confirmante	cortacorriente
chocante	confortante	cortante
cicatrizante	conglomerante	cosecante
ciente	conglutinante	cotangente
cimbreaente	congregante	covalente
circulante	conregnante	creciente
circunstante	consciente	crocante
circunyacente	consenciente	cuadrante
clarividente	conservante	cuadrupeante
cliente	consiguiente	cuentadante
coadyuvante	consintiente	culminante
coagente	consonante	culpante
coagulante	constante	cuociente
coalescente	constringente	curruscante
coamante	consultante	curtiente

D

danzante	descendente	diciente
debilitante	descendiente	diciplinante
debutante	descoagulante	dictante
decadente	descongestionante	difidente
decente	desconveniente	difrangente
deciente	desecante	dignificante
declarante	desemblante	diletante
declinante	desemejante	diligente
decolorante	desesperante	dirigente
decumbente	desgrasante	discante
defatigante	deshidratante	discente
deficiente	desincrustante	disciplinante
degenerante	desinfectante	disertante
degollante	desipiente	disgregante
degradante	deslumbrante	disidente
dehiscente	desobediente	disonante
dejante	desodorante	displicente
delante	desodorizante	distante
deliberante	desopilante	divergente
delicuescente	desoxidante	docente
delineante	desoxigenante	dodrante
demandante	despachante	dogmatizante
demulcente	despampanante	doliente
denante	despediente	dominante
denigrante	despeluznante	donante
denunciante	despidiente	dormiente
Deo volente	desplante	dotante
dependente	desternillante	dragante
dependiente	detergente	drogodependiente
deprecante	determinante	durante
derechohabiente	detonante	durmiente
desbordante	dibujante	

descacharrante dicente

E

ecuate	entremiente	estudiante
edulcorante	envergonzante	estupefaciente
efervescente	equipolente	euforizante
eficiente	equivalente	europaizante
eflorescente	errante	evacuante
ejecutante	erubesciente	evanescente
ejercitante	esbatimentante	evidente
embargante	escalofriante	examinante
emergente	escamante	excedente
emiente	esciente	excelente
emigrante	escribiente	excipiente
emoliente	espante	excitante
empeciente	espeluznante	excombatiente
emulgente	espinescente	exigente
enante	esquinante	exorbitante
encante	estante	expectante
engargante	estimulante	expectorante
enmiente	estresante	expediente
ensemejante	estridente	extravagante
enseñante	estuante	exuberante

F

fabricante	ferviente	formicante
faciente	figurante	fragante
farfante	fijante	fragante
farsante	filtrante	fraguante
fascinante	flagelante	frangente
febricitante	flagrante	frigente
fefaciente	flamante	fulgente
fehaciente	fletante	fulminante
fendiente	florecente	fumante

feriante	flotante	fundente
fertilizante	fluorescente	furiente

G

galiparlante	garante	gobernante
galopante	gestante	golfante
garamante	girante	gradiente

H

habiente	hendiente	hormigante
habitante	hiante	horripilante
hablante	hierofante	horrisonante
haciente	hilarante	hospedante
hebraizante	hispanohablante	humillante

I

ignorante	indehiscente	insolente
impaciente	independente	instante
impediente	independiente	instridente
impelente	indicante	insuficiente
imperante	indigente	insultante
impermeabilizante	indolente	insurgente
importante	indulgente	integrante
impresionante	inelegante	inteligente
improcedente	inerrante	intemperante
imprudente	infamante	intendente
impudente	infante	intercadente
incandescente	inficiente	interesante
incesante	infidente	interrogante
incidente	informante	interyacente
inciente	ingente	intolerante
incipiente	ingrediente	intransigente
incitante	inmigrante	intrascendente
inconducente	inmunizante	intrigante

inconfidente	innocente	intumescete
inconsciente	inobediente	invariante
inconsiguiente	inobservante	invidente
inconstante	inocente	iridiscente
inconveniente	inoperante	irrelevante
incrasante	insignificante	italianizante
incrustante	insipiente	itinerante

J

joyante	judicante	jurisprudente
judaizante	jurante	justificante

L

laborante	laxante	lubricante
labrante	lente	lubrificante
lactante	letificante	lugarteniente
lactescente	levante	lujurante
lampante	litigante	luminiscente
lancinante	lobagante	
latitante	lubigante	

M

machacante	mantiniente	ministrante
maestrante	marcescente	moderante
magnificente	marchante	modificante
maitinante	mareante	moliente
malandante	marinante	monovalente
maldiciente	marzante	montante
maleante	mediante	mordante
malevolente	medicinante	mordente
maloliente	mendicante	mordicante
malqueriente	mendigante	mordiente
malsonante	menguante	moviente
malviviente	menstruante	mulante

mandante	mercadante	munificente
mangante	mercante	musicante
manifestante	merchante	mutante
manteniente	militante	mutuante

N

naciente	negociante	nigromante
narcotraficante	némine discrepante	nocente
navegante	nesciente	nutriente
negligente	nictitante	

O

obediente	ocupante	operante
objetante	odorante	opinante
obsequiante	oficiante	ordenante
observante	olifante	oriente
obsolescente	omnisapiente	oscilante
obstante	omnisciente	otorgante
occidente	ondisonante	ovante
octante	opalescente	oxidante

P

paciente	piante	presidente
palente	picante	prestante
palpitante	pimpante	pretendiente
pariente	pinjante	previdente
parlante	pirante	principiante
participante	pivotante	procedente
parturiente	placiente	proficiente
pasante	plante	proliferante
pasavante	plurivalente	promesante
pasavolante	poderdante	prometiente
paseante	poderhabiente	protestante
pecante	polivalente	protuberante

pedante	poniente	providente
pediente	portante	proyectante
penante	posante	prudente
pendiente	postulante	pubescente
penetrante	practicante	pudiviente
pentavalente	precedente	pugnante
percuciente	precelente	pujante
perficiente	precipitante	pujavante
persevante	predicante	punzante
pesante	prefulgente	purgante
pescante	premoriente	putrescente
pestilente	preopinante	

Q

quemante	quiescente	quintante
----------	------------	-----------

R

radiante	refulgente	restante
rallante	regante	restaurante
rampante	regente	restringente
rapante	rehilante	resultante
rasante	reimplante	reticente
raspante	relajante	retumbante
reactante	relente	rigente
recalcitrante	relevante	rimbombante
reciente	reluctante	roborante
recipiente	rematante	rocinante
recitante	renunciante	rompiente
reclamante	repelente	rorante
recomendante	repitiente	rozagante
reconfortante	representante	ruante
recusante	repugnante	rubefaciente
redoblante	requirente	rubesciente
redolente	resfriante	rubricante

redoliente	residente	rumiante
reemplazante	resolvente	runflante
refriante	resonante	rusiente
refrigerante	resplendente	

S

saliente	significante	subdominante
salvante	siguiente	subiente
sangrante	silbante	subintendente
sargente	silente	subintrante
satirizante	simiente	subordinante
secante	simpatizante	subsiguiente
sedante	sirviente	subteniente
sedente	sobrante	subyacente
sedicente	sobrediente	sudante
sediente	sobresaliente	suficiente
semblante	sobrestante	superabundante
semejante	sobreviviente	superante
semiconsonante	solicitante	superdominante
semoviente	sonante	superintendente
senescente	sonriente	superviviente
serpiente	sorprendente	suplente
sextante	suavizante	suplicante
sibilante	subconsciente	sustentante
sicofante	subdelegante	

T

tajante	tintineante	trasplante
talante	tiplisonante	trasteante
talente	tirante	tratante
tangente	tocante	tremulante
televidente	tolerante	trepante
temblante	tomacorriente	trepidante
temperante	tonante	tributante

tenante	torneante	tridente
tendente	trabajante	trinchante
tendiente	trabucante	tripulante
teniente	trafagante	triumfante
teorizante	traficante	trivalente
terebrante	tragante	tronchante
terminante	trajinante	tunante
terrateniente	tranquilizante	tundente
tertuliante	transluciente	turbante
tetravalente	trasluciente	turgente

U

undante	urgente	urticante
---------	---------	-----------

V

vacante	vesicante	vigilante
vaciante	viajante	viniente
vagante	viandante	(ant. part. act. de venir)
valente	vibrante	visitante
valiente	vicariante	viviente
variante	vicealmirante	volante
veraneante	vicepresidente	volteante
vergonzante	vidente	vomipurgante
vertiente	vigente	votante

Y

yacente	yaciente	yente
---------	----------	-------

(de ir; lat. iens, euntis)

Anexo I. Diccionario de Datos

Table name	Table type	Primary key	# columns
ACCIONES	independent	ClaveVerbo	4
BASADAS_EN	dependent	Pregunta, ClaveObjRef	2
BASADAS_EN1	dependent	Pregunta, ClaveVerbo	2
CONTIENEN	dependent	Pregunta, ClaveEvento	2
EMPRESAS	independent	ClaveEmp	3
EVENTOS	independent	ClaveEvento	2
NOTICIAS	dependent	ClaveEmp, Fecha, Noticia	8
PREGUNTAS	independent	Pregunta	2
REFERENTES	independent	ClaveObjRef	2

Table 'ACCIONES'

Table name	ACCIONES
Primary key	pk_ACCIONES

Primary key 'pk_ACCIONES'

Column name	Type	Parent table
ClaveVerbo	PK	-----

Columns

Key	Column name	Domain	Data type	N ¹⁾	U ²⁾	C ³⁾	D ⁴⁾	Description
PK	ClaveVerbo		Integer	YES	NO	NO	NO	Clave autoincrementable de acciones
	NomVerbo		Char (20)	NO	NO	NO	NO	Descriptivo del verbo o acción
	Raiz		Char (20)	NO	NO	NO	NO	El verbo lematizado
	Tipo		Char (1)	NO	NO	NO	NO	Principal o Auxiliar

Relationships

Constraint name	Relationship type	Parent table	Child table	Card.
impican	Non-identifying	ACCIONES	NOTICIAS	1:N
	Identifying	ACCIONES	BASADAS_EN1	1:N

Description

Comprende los verbos principales a detectarse en la noticia

Table 'BASADAS_EN'

Table name	BASADAS_EN
Primary key	pk_BASADAS_EN

Primary key 'pk_BASADAS_EN'

Column name	Type	Parent table
Pregunta	PFK	PREGUNTAS
ClaveObjRef	PFK	REFERENTES

Columns

Key	Column name	Domain	Data type	N ¹⁾	U ²⁾	C ³⁾	D ⁴⁾	Description
PFK	Pregunta		Smallint	YES	NO	NO	NO	
PFK	ClaveObjRef		Integer	YES	NO	NO	NO	

Relationships

Constraint name	Relationship type	Parent table	Child table	Card.
	Identifying	PREGUNTAS	BASADAS_EN	1:N
	Identifying	REFERENTES	BASADAS_EN	1:N

Description

Registra los referentes y preguntas genéricas que permitirán las consultas generadas por los usuarios

Table 'BASADAS_EN1'

Table name	BASADAS_EN1
Primary key	pk_BASADAS_EN1

Primary key 'pk_BASADAS_EN1'

Column name	Type	Parent table
Pregunta	PFK	PREGUNTAS
ClaveVerbo	PFK	ACCIONES

Columns

Key	Column name	Domain	Data type	N ¹⁾	U ²⁾	C ³⁾	D ⁴⁾	Description
PFK	Pregunta		Smallint	YES	NO	NO	NO	
PFK	ClaveVerbo		Integer	YES	NO	NO	NO	

Relationships

Constraint name	Relationship type	Parent table	Child table	Card.
	Identifying	ACCIONES	BASADAS_EN1	1:N
	Identifying	PREGUNTAS	BASADAS_EN1	1:N

Description

Registra los verbos y preguntas genéricas que permitirán las consultas generadas por los usuarios

Table 'CONTIENEN'

Table name	CONTIENEN
Primary key	pk_CONTIENEN

Primary key 'pk_CONTIENEN'

Column name	Type	Parent table
Pregunta	PFK	PREGUNTAS
ClaveEvento	PFK	EVENTOS

Columns

Key	Column name	Domain	Data type	N ¹⁾	U ²⁾	C ³⁾	D ⁴⁾	Description
PFK	Pregunta		Smallint	YES	NO	NO	NO	
PFK	ClaveEvento		Smallint	YES	NO	NO	NO	

Relationships

Constraint name	Relationship type	Parent table	Child table	Card.
	Identifying	PREGUNTAS	CONTIENEN	1:N
	Identifying	EVENTOS	CONTIENEN	1:N

Description

Permite la posibilidad de que varios eventos se asignen a más de una pregunta

Table 'EMPRESAS'

Table name	EMPRESAS
Primary key	pk_EMPRESAS

Primary key 'pk_EMPRESAS'

Column name	Type	Parent table
ClaveEmp	PK	-----

Columns

Key	Column name	Domain	Data type	N ¹⁾	U ²⁾	C ³⁾	D ⁴⁾	Description
PK	ClaveEmp		Integer	YES	NO	NO	NO	Clave autoincrementable de empresa
	NombreEmp		Char (40)	NO	NO	NO	NO	Nombre de la empresa por revisar
	Liga		Varchar (120)	NO	NO	NO	NO	Página web donde se localizan las noticias

Relationships

Constraint name	Relationship type	Parent table	Child table	Card.
relativas_a	Identifying	EMPRESAS	NOTICIAS	1:N

Description

Maneja las empresas por procesar

Table 'EVENTOS'

Table name	EVENTOS
Primary key	pk_EVENTOS

Primary key 'pk_EVENTOS'

Column name	Type	Parent table
ClaveEvento	PK	-----

Columns

Key	Column name	Domain	Data type	N ¹⁾	U ²⁾	C ³⁾	D ⁴⁾	Description
PK	ClaveEvento		Smallint	YES	NO	NO	NO	Clave autoincrementable del evento
	NomEvento		Char (60)	NO	NO	NO	NO	Descriptivo del evento realizado

Relationships

Constraint name	Relationship type	Parent table	Child table	Card.
	Identifying	EVENTOS	CONTIENEN	1:N

Description

Enlista los hechos que comprende cada una de las preguntas genéricas

Table 'NOTICIAS'

Table name	NOTICIAS
Primary key	pk_NOTICIAS

Primary key 'pk_NOTICIAS'

Column name	Type	Parent table
ClaveEmp	PFK	EMPRESAS
Fecha	PK	-----
Noticia	PK	-----

Columns

Key	Column name	Domain	Data type	N ¹⁾	U ²⁾	C ³⁾	D ⁴⁾	Description
PFK	ClaveEmp		Integer	YES	NO	NO	NO	Clave de la empresa cuya noticia se registra
PK	Fecha		Date	YES	NO	NO	NO	Fecha de registro de la noticia
PK	Noticia		Smallint	YES	NO	NO	NO	Clave autoincrementable de noticia por empresa
	Agente		Varchar (60)	NO	NO	NO	NO	Descriptivo de la entidad que realiza la acción
FK	ClaveVerbo		Integer	YES	NO	NO	NO	Clave de la acción realizada
FK	ClaveObjRef		Integer	YES	NO	NO	NO	Clave del referente de objeto hallado
	ObjDet		Varchar (100)	NO	NO	NO	NO	Descriptivo del objeto detectado en la noticia
	Paciente		Varchar (60)	NO	NO	NO	NO	Descriptivo de quien recibe el efecto de la acción

Relationships

Constraint name	Relationship type	Parent table	Child table	Card.
implican	Non-identifying	ACCIONES	NOTICIAS	1:N
aparecen_en	Non-identifying	REFERENTES	NOTICIAS	1:N
relativas_a	Identifying	EMPRESAS	NOTICIAS	1:N

Description

Contiene la noticia en formato estructurado

Table 'PREGUNTAS'

Table name	PREGUNTAS
Primary key	pk_PREGUNTAS

Primary key 'pk_PREGUNTAS'

Column name	Type	Parent table
Pregunta	PK	-----

Columns

Key	Column name	Domain	Data type	N ¹⁾	U ²⁾	C ³⁾	D ⁴⁾	Description
PK	Pregunta		Smallint	YES	NO	NO	NO	Clave autoincrementable de la pregunta genérica
	DescPregunta		Varchar (100)	NO	NO	NO	NO	Texto descriptivo de la pregunta

Relationships

Constraint name	Relationship type	Parent table	Child table	Card.
	Identifying	PREGUNTAS	BASADAS_EN	1:N
	Identifying	PREGUNTAS	CONTIENEN	1:N
	Identifying	PREGUNTAS	BASADAS_EN1	1:N

Description

Engloba las preguntas a ser respondidas para fines de la inteligencia competitiva

Table 'REFERENTES'

Table name	REFERENTES
Primary key	pk_REFERENTES

Primary key 'pk_REFERENTES'

Column name	Type	Parent table
ClaveObjRef	PK	-----

Columns

Key	Column name	Domain	Data type	N ¹⁾	U ²⁾	C ³⁾	D ⁴⁾	Description
PK	ClaveObjRef		Integer	YES	NO	NO	NO	Clave autoincrementable de los referentes de objetos
	NomObjRef		Char (30)	NO	NO	NO	NO	Descripción del referente de Objeto

Relationships

Constraint name	Relationship type	Parent table	Child table	Card.
aparecen_en	Non-identifying	REFERENTES	NOTICIAS	1:N
	Identifying	REFERENTES	BASADAS_EN	1:N

Description

Mantiene los referentes de objetos contenidos en la noticia

¹⁾Not null

²⁾Unique

³⁾Check

⁴⁾Default

Column report

Column name	Table name	Data type	PK	FK
Agente	NOTICIAS	Varchar	NO	NO
ClaveEmp	EMPRESAS	Integer	YES	NO
ClaveEmp	NOTICIAS	Integer	YES	YES
ClaveEvento	EVENTOS	Smallint	YES	NO
ClaveEvento	CONTIENEN	Smallint	YES	YES
ClaveObjRef	REFERENTES	Integer	YES	NO
ClaveObjRef	BASADAS_EN	Integer	YES	YES
ClaveObjRef	NOTICIAS	Integer	NO	YES
ClaveVerbo	BASADAS_EN1	Integer	YES	YES
ClaveVerbo	ACCIONES	Integer	YES	NO
ClaveVerbo	NOTICIAS	Integer	NO	YES
DescPregunta	PREGUNTAS	Varchar	NO	NO
Fecha	NOTICIAS	Date	YES	NO
Liga	EMPRESAS	Varchar	NO	NO
NombreEmp	EMPRESAS	Char	NO	NO
NomEvento	EVENTOS	Char	NO	NO
NomObjRef	REFERENTES	Char	NO	NO
NomVerbo	ACCIONES	Char	NO	NO
Noticia	NOTICIAS	Smallint	YES	NO
ObjDet	NOTICIAS	Varchar	NO	NO
Paciente	NOTICIAS	Varchar	NO	NO
Pregunta	BASADAS_EN1	Smallint	YES	YES
Pregunta	CONTIENEN	Smallint	YES	YES
Pregunta	PREGUNTAS	Smallint	YES	NO
Pregunta	BASADAS_EN	Smallint	YES	YES
Raiz	ACCIONES	Char	NO	NO
Tipo	ACCIONES	Char	NO	NO

ANEXO J. Resultados del procesamiento de las noticias

empresa	fecha	agente	verbo	objeto ref.	objeto det.	paciente	estructura
UCC	01/03/2006	La Dirección General de Extensión Universitaria a través del Centro de l	haber, preparar	deportes y actividades recreativas			AVO
UCC	01/03/2006	La Universidad Cristóbal Colón	presentar	la convocatoria semestral	participar en programas de movilidad	los alumnos de licenciatura	AVO
UCC	01/03/2006	La Parroquia Universitaria	invitar	las celebraciones eucarísticas			AVO
UCC	01/03/2006	el rey Carlos III	encomendar	la traducción	la Biblia completa al castellano		AVO
UCC	01/03/2006		Acreditar	Certificación Microsoft Office Specialist nivel Core			VO
UCC	01/03/2006	Padres Escolapios radicados en Bogotá, Colombia	contar	una nueva colección	microfilmaciones	Dr. Segismundo Balagué S.", d	PVOA
UCC	01/03/2006	el Instituto Veracruzano de la Cultura (IVEC)	entregar	Donación	acervo bibliográfico	la Cátedra en Comunicación y	AVOP
UCC	01/03/2006	El ICCE: Educación Continua	iniciar	el Servicio	Estimulación y Rehabilitación Psicopedagógica		AVO
UCC	01/03/2006	de la licenciatura en Ciencias de la Comunicación de la Universidad Cris	producir, transmitir	programas			AVO
UCC	01/03/2006	El Consejo Mexicano de Acreditación de Enseñanza de la Arquitectura, .	acreditar	acreditación del COMAEA		la licenciatura en Arquitectura e	AVP
UCC	01/03/2006	el edificio	obtener	mención de honor		Residencia universitaria del car	APVO
UCC	01/03/2006	la comunidad educativa	informar	las nuevas disposiciones	bibliotecas		VPO
UCC	01/03/2006	Senectos y Activos (hoy Adultos Mayores)	acudir, realizar	convenio	licenciatura en Ciencias de la Comunicación		AVO
LA SALLE	01/03/2006	Fidelidad y servicio	ser				AVO
LA SALLE	01/03/2006	los programas de Educación Continua de la Universidad De La Salle Ba	ofrecer	diplomado	Cocina Mexicana a través de los Siglos		AVO
LA SALLE	01/03/2006	Nuestra Universidad De La Salle Bajío	er, engalanar y presenta	el concierto	música italiana y mexicana		AVO
LA SALLE	01/03/2006	la Comunidad Lasallista de la ciudad	realizar	Celebración Eucarística	Centenario Lasallista		OVA
LA SALLE	01/03/2006	La Biblioteca de la Universidad De La Salle Bajío	ser				AVO
LA SALLE	01/03/2006	Diseño ambiental	ser				OVA
LA SALLE	01/03/2006	La Universidad De La Salle Bajío	iniciar				AVO
LA SALLE	01/03/2006	Ing. Julio Di Bella Roldán, Director General del Canal Once de televisión,	llevar a cabo	la conferencia	La empresa y los medios de comunicación		VOA
LA SALLE	01/03/2006	el Padre Rubén Porras	ofrecer	la Celebración Eucarística			AVO
LA SALLE	01/03/2006	Universidad De La Salle Bajío	llevar a cabo	el evento	Dame una U en este Campus Salamanca		AVO
LA SALLE	01/03/2006	al Odontólogo	temer	Día		La mayoría de los pacientes	PVA
LA SALLE	01/03/2006	Los Hermanos	tributar	el homenaje que le deben	como su fundador	San Juan Bautista De La Salle	AVPO
LA SALLE	01/03/2006	El Centro Nacional de Evaluación para la Educación Superior A.C. CEN	preparar	Examen General de Cond	para los egresados de la Licenciatura en Derecho		AVO
LA SALLE	01/03/2006	nuestra Universidad De La Salle Bajío	asistir	una Plática de Información			AVO
LA SALLE	01/03/2006	El Centro de Lenguas de nuestra Universidad	albergar				AVO
LA SALLE	01/03/2006	La Escuela de Veterinaria	ber, desarrollar y fortalecer				AV
LA SALLE	01/03/2006	el Dr. Román Esqueda Ataide, Director del despacho Posicionamiento y	versar	el curso de Capacitación	en técnicas persuasivas para los diseños		VOA
UDLA	01/03/2006	La UDLA	llevar a cabo	XVI Congreso	Ciencias de la Comunicación		AVO
UDLA	01/03/2006	UDLA		campeón	estatal del CONDDE		AO
UDLA	01/03/2006	en la UDLA	inagurar	XVI Congreso	Internacional de Electrónica, Comunicaciones y Computadoras		VAO
UDLA	01/03/2006	El Cine Club de la UDLA	proyectar	la película	"Perdón María"		AVO
UDLA	01/03/2006	Canal 6 de julio	presentar	más reciente producción	en la UDLA		AVO
UDLA	01/03/2006	La Compañía de Teatro de la Universidad de las Américas Puebla (UDL	presentar	la obra	"Las Preciosas Ridículas"		AVO
UDLA	01/03/2006	Iraida Noriega	deleitar	Temporada	Cultural de la UDLA		AVO
UDLA	01/03/2006	UDLA-Angelópolis	poner, bailar			público poblano	AVP
UNITEC	01/03/2006	UNITEC	inaugurar	Segunda Etapa	Campus Coyoacán		VAO
UNITEC	01/03/2006	UNITECC y OCC	firmar	Convenio de Colaboración			AVO

UNITEC	01/03/2006	La SEP	otorgar	certificación	"Excelencia Académica"		AVO
UNITEC	01/03/2006	UNITEC y Adecco	firmar	Convenio de Intercambio			AVO
UNITEC	01/03/2006	La incubadora de Empresas en Negocios Electrónicos y Tecnología de I	recibir	importantes reconocimientos			AVO
UNITEC	01/03/2006	UNITEC y la Asociación de Empresarios de Atizapán	firmar	Convenio de Colaboración			VOA
UNITEC	01/03/2006	UNITEC y Manpower	firmar	Convenio de Colaboración			AVO
UNITEC	01/03/2006	El DR. Takao Kimura, director de programa de posgrado de odontología	recibir	el premio	Master for Excellence in Health Care (maestro por la excelencia en el cuidado de la salud)		AVO
UNITEC	01/03/2006			Certificación	Proveedor Autorizado de la IACET	División de Posgrado de la UNI	OP
UNITEC	01/03/2006			premio	Especial Obras Cemex de Arquitectura Sustentable	Campus Zapopan, de la UNITE	OP
UPAEP	01/03/2006	los alumnos de Administración de Instituciones de la Universidad Popula	retornar	casa de estudios			AVO
UPAEP	01/03/2006	UPAEP	calificar	regional	Fútbol Rápido		AVO
UPAEP	01/03/2006	La artista Pepita Albisúa	donar	obra completa		al museo	AVOP
UPAEP	01/03/2006	Televisa	ofrecer	plática	alumnos de la UPAEP para asistir a Espacio 2006		AVO
UPAEP	01/03/2006	el CACEI	Acreditar			Ingeniería Mecatrónica de la U	PVA
UPAEP	01/03/2006	UPAEP	lograr, clasificar	fase de semifinales	Universidad Nacional que organiza el consejo nacional del deporte de la educacion conada		AVO
UPAEP	01/03/2006	UPAEP	retomar			la UMAD	AVP
UPAEP	01/03/2006	UPAEP	conseguir	victoria	Centro de Estudios Universitarios de Monterrey en la Conferencia Nacional de Baloncesto		AVO
UDEM	01/03/2006	las comisiones	discutir		El proceso de vinculación preparatoria-profesional, el perfil del profesor UDEM, las necesida		OAV
UDEM	01/03/2006	Karina Astorga Carrasco	asegurar	Centro de Incubación y D	Promover el auto empleo, la innovación de productos y cre	es el principal objetivo	OVA
UDEM	01/03/2006	la División de Ciencias de la Salud (DICSA)	elaborar	altar de muertos		en la	AVO
UDEM	01/03/2006	la comunidad universitaria	poder, visitar	los cuartos			AVO
UDEM	01/03/2006		ser, enmarcar	la celebración	36 aniversario de la UDEM		VO
UDEM	01/03/2006	El municipio de san pedro	realizar	recorridos	centros nocturnos los fines de semana		AVO
UDEM	01/03/2006		ser	las aportaciones	dejaron las asociaciones		VO
UDEM	01/03/2006	alumnos y personal de la institución	unir	una cadena de oración	dar gracias por un año mas de servir y formar personas comprometidas con su comunidad		AVO
UDEM	01/03/2006	200 alumnos de Ingles, Frances, Aleman y Español	quedar sin	actividades	laboratorio debido a un problema dentro del software		AVO
UDEM	01/03/2006	los ex alumnos egresados de la UDEM	poder comprar	ideales profesionales, personales y familiares de ese entonces			AVO
UDEM	01/03/2006	Julieta Alejandra Yañez Vega	fijar	Apertura, intercambios estudiantiles e internacionalización			OVA
UDEM	01/03/2006	ivisión de Arquitectura, Ingeniería y Diseño en asociación con Dynatek	tener	mejor acercamiento	su carrera con cursos como el impartido	de la Ingeniería en Mecatrónica	PVOA
UDEM	01/03/2006	la ex alumna de la lic. En artes, Estela Torres	quedar plamado	El símbolo de humanismo			OVA
UDEM	01/03/2006	el centro de orientación estudiantil(COES)	implementar	nuevo programa	tutoría de pares		OVA
UDEM	01/03/2006	Oscar Flores Torres	estar-dar-conocer	la conferencia	Vida y Obra de Alberto Pani		OVA
UDEM	01/03/2006	Roberto Adame Doria, director de ADC Consultores de Negocios	señalar	La formación	un emprendedor eficiente comienza en la familia, estudios en segundo lugar y las empresa		OVA
UDEM	01/03/2006	Gustavo Raymundo Gallo y Eduardo Suárez Fresnillo en Comunicando	ser	la observación	¿Foráneos? En sus ciudades los necesitan, necesitan profesionistas, en Monterrey ya no s		OVA
UDEM	01/03/2006	Francisco José Navarro Aguirre	impartir	Virus del Papiloma Humano, enfermedad número uno en Nuevo León	a los alumnos de la UDEM		POAV
UDEM	01/03/2006	Los alumnos de Derecho	contar	un Aula	Destrezas Jurídicas		AVO
UDEM	01/03/2006	Abraham Daniel González Gloria	instar - proteger	la propiedad intelectual.			AVO
UDEM	01/03/2006	La premiación Repentina 2005	seleccionar	cinco trabajos	ganadores del concurso interno de diseño arquitectónico		AVO
UDEM	01/03/2006	Un total de 14 personas	tramitar	credencial	el uso del elevador		AVO
UDEM	01/03/2006	a la UDEM,	poder - solucionar	una planta de emergencia			AVO
UDEM	01/03/2006		ser-reconocer	Premio Nacional de la Federación 2005 a la Excelencia Educativa		el maestro de Ingeniería y Educ	VOA
UDEM	01/03/2006		ser	panel Juventud Empresar	Ser diferentes en el mercado, obtener una mentalidad soñadora, tomar el riesgo, el cambio		OV
UDEM	01/03/2006	María de los Ángeles Pozas	ser-presentar	libro	Estructura y Dinámica de la Gran Empresa en México		OVA
UDEM	01/03/2006	María de la Luz Bravo Santillán	acordar	su conferencia	Balance Scorecard: herramienta para Implementar Estrategias.		VAO
UDEM	01/03/2006	la EBC	realizar	una convivencia	los alumnos	Bachillerato y Licenciatura	AVO

UDEM	01/03/2006	Patricia Mercado Partido de Alternativa Socialdemócrata	dar - conocer	plataforma política			AVOP
UDEM	01/03/2006	Tres instituciones	invertir	millones de pesos			AVO
UDEM	01/03/2006	la Escuela Bancaria y Comercial	establecer	convenios e intercambios	con universidades del extranjero		AVO
UDEM	01/03/2006	La EBC	ofrecer	opción educativa y hoteles	para gente de negocios		AVO
UDEM	01/03/2006	el Presidente del Colegio	mencionar	la importancia			AVO
UDEM	01/03/2006	factores económicos	dejar- estudiar	dos terceras partes de los jóvenes			AVO
ITESO	01/03/2006		estar-invitar-asistir	Cuarto Informe del Rector		La comunidad universitaria	PVO
ITESO	01/03/2006	Café Scientifique Guadalajara	invitar	segunda charla del año			AVO
ITESO	01/03/2006	Manuel Ramírez, académico del ITESO	decir	la presentación	los tres tomos del libro		VAO
ITESO	01/03/2006	Hector Toledano	concebir	la vocación	de escritor		VAO
ITESO	01/03/2006	ITESO					AVO
ITESO	01/03/2006	Gabriel Morales académico del Instituto de Formación Filosófica Interco	obedecer-decir	plática-proyectos neoliberales	"El claroscuro de la formación en competencias profesionales"		VAO
ITESO	01/03/2006	la revista Sinéctica	hablar	la conformación	del oficio docente		AVO
ITESO	01/03/2006	Fernando Jiménez del Consejo de Cámaras Industriales de Jalisco (CCI	estar acordar	La situación	los empresarios en México		OVA
ITESO	01/03/2006	ITESO, Ciatej y UdeG Biocluster de Occidente	pretender	La creación de una entidad			AVO
ITESO	01/03/2006		llevar a cabo	el panel	"De estudiante a empleador, ¡emprende ya!"		VO
ITESO	01/03/2006	Anita Nielsen Dhont, profesora del Departamento de Educación y Valores	haber-enfocar-señalar		"El aprendizaje en un escenario de experiencias complejo, que requiere reto al pensamiento"		AVO
ITESO	01/03/2006	Tipo Polo	dar forma-producir	documental Arcediano			VO
ITESO	01/03/2006	Universidad Iberoamericana campus Puebla	realizar		"Mexico: las encrucijadas de la democracia"		OVA
ITESO	01/03/2006	Alumnos de Diseño	voltear	al mundo prehispánico			AVO
ITESO	01/03/2006	El inglés	necesitar-acceder	mundo laboral			OVA
ITESO	01/03/2006	Israel Dehesa - Casa de Estudios	informar		"La Primera Muestra de Cortos Mexicanos"		OVA
ITESO	01/03/2006	el Centro de Promoción Cultural	ofrecer	el ciclo Cineforum			OAV
ITESO	01/03/2006	la actriz y dramaturga Circee Rangel	opinar-impartir	del taller	Teatro en Técnicas de la Improvisación		AOV
ITESO	01/03/2006	El ITESO, la Universidad Jesuita en Guadalajara	ser	La institución de educación	importante de la región occidente		AVO
CETYS	01/03/2006	Los Zorros y las Zorritas de CETYS Universidad	continuar	racha ganadora			AVO
CETYS	01/03/2006	Instituciones de Educación Superior	visitar	las instalaciones	CETYS Universidad		AVO
CETYS	01/03/2006	El Ing. Enrique Carlos Blancas de la Cruz, Rector del Sistema de CETYS	ser	Reunión Semanal	la Asociación de Periodistas de Mexicali (APM)		AVO
CETYS	01/03/2006	los estudiantes de Preparatoria de CETYS Universidad	llevar a cabo	a presentación de comerciales			VOA
CETYS	01/03/2006	los Egresados de CETYS Universidad, Iliana Coria, Juan José Esquivias	decorar				AV
CETYS	01/03/2006	Los equipos de CETYS Universidad	ganar	encuentros			AVO
CETYS	01/03/2006	El Gobierno Mexicano	destinar	dos mil millones de pesos		pequeñas y medianas empresas	AVO
CETYS	01/03/2006	Bancomext	destinar			a las pequeñas y medianas empresas	AVP
CETYS	01/03/2006		iniciar	los nuevos cursos	la Alianza Francesa		VOP
CETYS	01/03/2006	Egresado del CETYS	contratar-platicar			los estudiantes	AVP
CETYS	01/03/2006		realizar	Conferencia	la India y el Hinduismo		AV
CETYS	01/03/2006	Lic. Javier Oliver Gallego de Mercadotecnia de CETYS Universidad	impartir - tener	una conferencia	"Experiencias de un Emprendedor"	de Administración de Empresas	AVO
CETYS	01/03/2006	CETYS Universidad	recibir	Congreso de	Gerontología Social		OAV
ANAHUAC	01/03/2006	la Universidad Anáhuac y el INAH	firmar	convenio			VOA
ANAHUAC	01/03/2006	Miembros de CREA Jim Caviezel y José Bastón	convivir				AV
ANAHUAC	01/03/2006			Medalla Anáhuac	en Diseño	Carmen Cordera y Giuseppe Zecchi	OP
ANAHUAC	01/03/2006	Nuestra Universidad	ser	Encuentro	Iberoamericano de Protección de Datos Personales		AVO
ANAHUAC	01/03/2006	Alumnos del Lic. Pablo González Cid	convivir				AV
ANAHUAC	01/03/2006	La Escuela de Diseño de la Universidad Anáhuac	realizar	Congreso Internacional	"Diseñar para la Humanidad"		AVO

ANAHUAC	01/03/2006	Dr. Carlos M. Jarque Uribe	<i>insistir</i>	<i>Doctor Honoris Causa</i>			VOP
ANAHUAC	01/03/2006	Acción Social de la Universidad Anáhuac (ASUA)	<i>organizar</i>	<i>importantes eventos</i>			AVO
ANAHUAC	01/03/2006	Profesores del Centro de Negocios Agropecuarios Internacionales (CEN	<i>recibir</i>	<i>la acreditación</i>	<i>Financiera Rural</i>		PVO

ANEXO K Resultados del procesamiento de las noticias (Con sustitución)

empresa	fecha	agente	verbo	objeto ref.	objeto det.	paciente	estructura
2	01/03/2006	La Dirección General de Extensión Universitaria a través del Centro d	111	54			AVO
2	01/03/2006	La Universidad Cristóbal Colón	67	48	participar en programas de movilidad	los alumnos de licenciatura	AVO
2	01/03/2006	La Parroquia Universitaria	52	26			AVO
2	01/03/2006	el rey Carlos III	98	151	la Biblia completa al castellano		AVO
2	01/03/2006		88	30			VO
2	01/03/2006	Padres Escolapios radicados en Bogotá, Colombia	23	34	microfilmaciones	Dr. Segismundo Balagué S.", del campus Torrente Viver	PVOA
2	01/03/2006	el Instituto Veracruzano de la Cultura (IVEC)	38	63	acervo bibliográfico	la Cátedra en Comunicación y Estudios Culturales de la Un	AVOP
2	01/03/2006	El ICCE: Educación Continua	49	138	Estimulación y Rehabilitación Psicopedagógica		AVO
2	01/03/2006	La licenciatura en Ciencias de la Comunicación de la Universidad Cris	69	123			AVO
2	01/03/2006	El Consejo Mexicano de Acreditación de Enseñanza de la Arquitectura	88	3		la licenciatura en Arquitectura de la Universidad Cristóbal C	AVP
2	01/03/2006	el edificio	57	96		Residencia universitaria del campus Calasanz de la Universi	APVO
2	01/03/2006	la comunidad educativa	48	60	bibliotecas		VPO
2	01/03/2006	Senectos y Activos (hoy Adultos Mayores)	71	45	licenciatura en Ciencias de la Comunicación		AVO
17	01/03/2006	Fidelidad y servicio	78	0			AVO
17	01/03/2006	los programas de Educación Continua de la Universidad De La Salle	101	59	Cocina Mexicana a través de los Siglos		AVO
17	01/03/2006	Nuestra Universidad De La Salle Bajío	67	38	música italiana y mexicana		AVO
17	01/03/2006	la Comunidad Lasallista de la ciudad	71	26	Centenario Lasallista		OVA
17	01/03/2006	La Biblioteca de la Universidad De La Salle Bajío	78	0			AVO
17	01/03/2006	Diseño ambiental	78	0			OVA
17	01/03/2006	La Universidad De La Salle Bajío	49	0			AVO
17	01/03/2006	Ing. Julio Di Bella Roldán, Director General del Canal Once de televisi	71	41	La empresa y los medios de comunicación		VOA
17	01/03/2006	el Padre Rubén Porras	101	26			AVO
17	01/03/2006	Universidad De La Salle Bajío	71	69	Dame una U en este Campus Salamanca		AVO
17	01/03/2006	al Odontólogo	102	57		La mayoría de los pacientes	PVA
17	01/03/2006	Los Hermanos	82	82	como su fundador	San Juan Bautista De La Salle	AVPO
17	01/03/2006	El Centro Nacional de Evaluación para la Educación Superior A.C. Cl	103	70	para los egresados de la Licenciatura en Derecho		AVO
17	01/03/2006	nuestra Universidad De La Salle Bajío	90	116			AVO
17	01/03/2006	El Centro de Lenguas de nuestra Universidad	3	0			AVO
17	01/03/2006	La Escuela de Veterinaria	112	0			AV
17	01/03/2006	el Dr. Román Esqueda Ataide, Director del despacho Posicionamientc	85	53	en técnicas persuasivas para los diseños		VOA
1	01/03/2006	La UDLA	71	43	Ciencias de la Comunicación		AVO
1	01/03/2006	UDLA	0	20	estatal del CONDDE		AO
1	01/03/2006	en la UDLA	47	43	Internacional de Electrónica, Comunicaciones y Computadoras		VAO
1	01/03/2006	El Cine Club de la UDLA	70	109	"Perdón María"		AVO
1	01/03/2006	Canal 6 de julio	67	122	en la UDLA		AVO
1	01/03/2006	La Compañía de Teatro de la Universidad de las Américas Puebla (U	67	101	"Las Preciosas Ridículas"		AVO
1	01/03/2006	Iraida Noriega	95	146	Cultural de la UDLA		AVO
1	01/03/2006	UDLA–Angelópolis	113	0		público poblano	AVP
8	01/03/2006	UNITEC	47	68	Campus Coyoacán		VAO
8	01/03/2006	UNITECC y OCC	43	45			AVO

8	01/03/2006	La SEP	61	30	"Excelencia Académica"		AVO
8	01/03/2006	UNITEC y Adecco	43	45			AVO
8	01/03/2006	La incubadora de Empresas en Negocios Electrónicos y Tecnología d	72	130			AVO
8	01/03/2006	UNITEC y la Asociación de Empresarios de Atizapán	43	45			VOA
8	01/03/2006	UNITEC y Manpower	43	45			AVO
8	01/03/2006	El DR. Takao Kimura, director de programa de posgrado de odontolog	72	119	Master for Excellence in Health Care (maestro por la excelencia en el cuidado de la salud)		AVO
8	01/03/2006		0	30	Proveedor Autorizado de la IACET	División de Posgrado de la UNITEC	OP
8	01/03/2006		0	119	Especial Obras Cemex de Arquitectura Sustentable	Campus Zapopan, de la UNITEC	OP
10	01/03/2006	los alumnos de Administración de Instituciones de la Universidad Popu	75	24			AVO
10	01/03/2006	UPAEP	91	132	Fútbol Rápido		AVO
10	01/03/2006	La artista Pepita Albisúa	33	101		al museo	AVOP
10	01/03/2006	Televisa	101	116	alumnos de la UPAEP para asistir a Espacio 2006		AVO
10	01/03/2006	el CACEI	88	0		Ingeniería Mecatrónica de la UPAEP	PVA
10	01/03/2006	UPAEP	15	73	Universidad Nacional que organiza el consejo nacional del deporte de la educacion conadde		AVO
10	01/03/2006	UPAEP	104	0		la UMAD	AVP
10	01/03/2006	UPAEP	21	152	Centro de Estudios Universitarios de Monterrey en la Conferencia Nacional de Baloncesto de la CONADEIP		AVO
19	01/03/2006	las comisiones	96	0	El proceso de vinculación preparatoria-profesional, el perfil del profesor UDEM, las necesidades del aula como recurso		OAV
19	01/03/2006	Karina Astorga Carrasco	89	28	Promover el auto empleo, la innovación de productos y cre	es el principal objetivo	OVA
19	01/03/2006	la División de Ciencias de la Salud (DICSA)	97	9		en la	AVO
19	01/03/2006	la comunidad universitaria	86	52			AVO
19	01/03/2006		114	26	36 aniversario de la UDEM		VO
19	01/03/2006	El municipio de san pedro	71	131	centros nocturnos los fines de semana		AVO
19	01/03/2006		78	12	dejaron las asociaciones		VO
19	01/03/2006	alumnos y personal de la institución	83	18	dar gracias por un año mas de servir y formar personas comprometidas con su comunidad		AVO
19	01/03/2006	200 alumnos de Ingles, Frances,Aleman y Español	121	4	laboratorio debido a un problema dentro del software		AVO
19	01/03/2006	los ex alumnos egresados de la UDEM	115	83			AVO
19	01/03/2006	Julieta Alejandra Yañez Vega	105	11			OVA
19	01/03/2006	ivisión de Arquitectura, Ingeniería y Diseño en asociación con Dynatel	106	157	su carrera con cursos como el impartido	de la Ingeniería en Mecatrónica	PVOA
19	01/03/2006	la ex alumna de la lic. En artes, Estela Torres	116	139			OVA
19	01/03/2006	el centro de orientación estudiantil(COES)	99	123	tutoría de pares		OVA
19	01/03/2006	Oscar Flores Torres	48	41	Vida y Obra de Alberto Pani		OVA
19	01/03/2006	Roberto Adame Doria, director de ADC Consultores de Negocios	107	42	un emprendedor eficiente comienza en la familia, estudios en segundo lugar y las empresas incubadoras después		OVA
19	01/03/2006	Gustavo Raymundo Gallo y Eduardo Suárez Fresnillo en Comunicand	78	102	¿Foráneos? En sus ciudades los necesitan, necesitan profesionistas, en Monterrey ya no se necesitan, aquí ya hay de		OVA
19	01/03/2006	Francisco José Navarro Aguirre	46	155		a los alumnos de la UDEM	POAV
19	01/03/2006	Los alumnos de Derecho	23	15	Destrezas Jurídicas		AVO
19	01/03/2006	Abraham Daniel González Gloria	50	124			AVO
19	01/03/2006	La premiación Repentina 2005	77	150	ganadores del concurso interno de diseño arquitectónico		AVO
19	01/03/2006	Un total de 14 personas	81	50	el uso del elevador		AVO
19	01/03/2006	a la UDEM,	117	114			AVO
19	01/03/2006	el maestro de Ingeniería y Educación José Benito Flores Juárez	73	119			VOA
19	01/03/2006		78	108	Ser diferentes en el mercado, obtener una mentalidad soñadora, tomar el riesgo, el cambio de paradigmas, invertir en ca		OV
19	01/03/2006	María de los Ángeles Pozas	67	92	Estructura y Dinámica de la Gran Empresa en México		OVA
19	01/03/2006	María de la Luz Bravo Santillán	87	41	Balance Scorecard: herramienta para Implementar Estrategias.		VAO
19	01/03/2006	la EBC	71	47	los alumnos	Bachillerato y Licenciatura	AVO

19	01/03/2006	Patricia Mercado Partido de Alternativa Socialdemócrata	48	115			AVOP
19	01/03/2006	Tres instituciones	100	111			AVO
19	01/03/2006	la Escuela Bancaria y Comercial	39	45	con universidades del extranjero		AVO
19	01/03/2006	La EBC	101	105	para gente de negocios		AVO
19	01/03/2006	el Presidente del Colegio	108	85			AVO
19	01/03/2006	factores económicos	118	67	dos terceras partes de los jóvenes		AVO
15	01/03/2006	La comunidad universitaria	52	87			PVO
15	01/03/2006	Café Scientifique Guadalajara	52	31			AVO
15	01/03/2006	Manuel Ramirez, académico del ITESO	93	120	los tres tomos del libro		VAO
15	01/03/2006	Hector Toledano	92	156	de escritor		VAO
15	01/03/2006	Gabriel Morales académico del Instituto de Formación Filosófica Inter	93	116	"El claroscuro de la formación en competencias profesionales"		VAO
15	01/03/2006	la revista Sinéctica	45	42	del oficio docente		AVO
15	01/03/2006	Fernando Jiménez del Consejo de Cámaras Industriales de Jalisco (C	87	142	los empresarios en México		OVA
15	01/03/2006	ITESO, Ciatej y UdeG Biocluster de Occidente	68	49			AVO- OVA
15	01/03/2006		71	108	"De estudiante a empleador, ¡emprende ya!"		VO
15	01/03/2006	Anita Nielsen Dhont, profesora del Departamento de Educación y Valc	107	0	"El aprendizaje en un escenario de experiencias complejo, que requiere reto al pensamiento y a la mente, es donde el c		AVO
15	01/03/2006	Tipo Polo	119	62			VO
15	01/03/2006	Universidad Iberoamericana campus Puebla	71	0	"Mexico: las encrucijadas de la democracia"		OVA
15	01/03/2006	Alumnos de Diseño	109	98			AVO
15	01/03/2006	El inglés	120	98			OVA
15	01/03/2006	Israel Dehesa - Casa de Estudios	48	0	"La Primera Muestra de Cortos Mexicanos		OVA
15	01/03/2006	el Centro de Promoción Cultural	101	32			OAV
15	01/03/2006	la actriz y dramaturga Circee Rangel	46	143	Teatro en Técnicas de la Improvisación		AOV
15	01/03/2006	El ITESO, la Universidad Jesuita en Guadalajara	78	90	importante de la región occidente		AVO
13	01/03/2006	Los Zorros y las Zorritas de CETYS Universidad	25	127			AVO
13	01/03/2006	Instituciones de Educación Superior	86	89	CETYS Universidad		AVO
13	01/03/2006	El Ing. Enrique Carlos Blancas de la Cruz, Rector del Sistema de CET	78	135	la Asociación de Periodistas de Mexicali (APM)		AVO
13	01/03/2006	los estudiantes de Preparatoria de CETYS Universidad	71	120			VOA
13	01/03/2006	os Egresados de CETYS Universidad, Iliana Coria, Juan José Esquivi	94	0			AV
13	01/03/2006	Los equipos de CETYS Universidad	44	64			AVO
13	01/03/2006	El Gobierno Mexicano	31	111		pequeñas y medianas empresas (pymes) en 2006	AVO
13	01/03/2006	Bancomext	31	0		a las pequeñas y medianas empresas	AVP
13	01/03/2006		49	53	la Alianza Francesa		VOP
13	01/03/2006	Egresado del CETYS	64	0		los estudiantes	AVP
13	01/03/2006		71	41	la India y el Hinduismo		AV
13	01/03/2006	Lic. Javier Oliver Gallego de Mercadotecnia de CETYS Universidad	46	41	"Experiencias de un Emprendedor"	alumnos de Administración de Empresas	AVO
13	01/03/2006	CETYS Universidad	72	43	Gerontología Social		OAV
16	01/03/2006	la Universidad Anáhuac y el INAH	43	45			VOA
16	01/03/2006	Miembros de CREA Jim Caviezel y José Bastón	27	0			AV
16	01/03/2006		0	95	en Diseño	Carmen Cordera y Giuseppe Zecca	OP
16	01/03/2006	Nuestra Universidad	78	64	Iberoamericano de Protección de Datos Personales		AVO
16	01/03/2006	Alumnos del Lic. Pablo González Cid	27	0			AV
16	01/03/2006	La Escuela de Diseño de la Universidad Anáhuac	71	43	"Diseñar para la Humanidad"		AVO
16	01/03/2006	Dr. Carlos M. Jarque Uribe	110	158			VOP

16	01/03/2006	Acción Social de la Universidad Anáhuac (ASUA)	60	69			AVO
16	01/03/2006	Profesores del Centro de Negocios Agropecuarios Internacionales (CENAI)	72	3	Financiera Rural		PVO

Anexo L. Reglas para la elección del verbo principal

Votos a favor.

- ¿ La palabra anterior es un Nombre propio?
- ¿ Precede a un Objeto?
- ¿ Va después de la palabra "se"?
- ¿ Va después de verbo auxiliar, como haber?
- ¿ Va luego de la palabra "en" más una palabra intermedia?
- ¿ Continúa a un texto aclaratorio delimitado por una coma?

Votos en contra.

- ¿ Cuando el verbo está en participio y le antecede un nombre propio?
- ¿ Cuando el verbo está en participio y le antecede la palabra "que"?
- ¿ Cuando el verbo termina en "ión"/"iones"?
- ¿ Cuando el verbo termina en "an"?
- ¿ Cuando el verbo termina en "dad"?
- ¿ Cuando le antecede algún artículo determinativo ("el la los las al del")?
- ¿ Cuando está en infinitivo y le precede "de"?
- ¿ Cuando está en gerundio y le precede un verbo en infinitivo?