

Las tecnologías del lenguaje en el manejo de información y en la generación del conocimiento

GERARDO SIERRA MARTÍNEZ
*Instituto de Ingeniería,
Universidad Nacional Autónoma de México*

EL MANEJO DE INFORMACIÓN

En nuestra sociedad actual, la información y el conocimiento son componentes esenciales para el progreso económico y para el bienestar personal. A fin de lograr la generación de conocimiento para tener las respuestas más adecuadas ante las diversas circunstancias que se van presentando, las personas necesitan disponer de información oportuna y verídica.

Con la creación de una política inteligente de acceso a la información, se logra un avance importante para una interacción eficiente entre los sectores industriales y estatales —con necesidades concretas de conocimiento—, los sectores académicos —productores y gestores del mismo— y la sociedad. Al establecer un sistema de procesamiento eficiente de intercambio de conocimiento entre las partes mencionadas, se pueden realizar acciones concretas con resultados óptimos que ayuden realmente a los diferentes estratos de la población, por ejemplo pequeñas y medianas empresas (PyMEs) con objetivos de comercio internacional que coadyuven al desarrollo económico y sustentable del país.

Existe, en México, una gama considerable de empresas que, de una u otra forma, tienen que tratar con la gestión de información. Tan

sólo en el D.F., hay alrededor de 116,000 empresas registradas¹ que requieren de un tratamiento a mediana o gran escala para actualizar su información y gestionarla de manera que puedan tener una mejor competencia en el mercado y logren generar un mayor número de divisas. Igualmente, los institutos y entidades requieren de una gestión y una explotación de calidad de información eficiente, por lo que es menester tomarlos en cuenta para que México siga en su ruta de avance y progreso.

INFORMACIÓN, CONOCIMIENTO Y LENGUAJE

Si bien la información se ha tomado muchas veces como equivalente de saber o de conocimiento, hay varias diferencias entre estos dos conceptos que debemos considerar. Entre las características principales de la información podemos mencionar las siguientes: se trata de un bien de naturaleza intelectual, inmaterial, aunque se registra y presenta en soportes físicos. No es escasa; cada vez es más abundante y su uso no es excluyente. La pueden utilizar varias personas a la vez, salvo cuando su valor está precisamente en la privacidad y exclusividad. Por otro lado, respecto al conocimiento, decimos que éste es fruto de las asociaciones que el cerebro es capaz de hacer a partir del análisis de información almacenada. Además, el conocimiento se elabora para dar respuesta a alguna cuestión de interés particular; por tanto, el conocimiento siempre supone más que la información de la que se ha partido para construirlo; tiene una funcionalidad.

Pese a sus diferencias, existe una relación indisoluble entre información, conocimiento y lenguaje; tanto la información como el conocimiento se expresan y transmiten gracias al lenguaje (oral, escrito, gesticular, etc.) La “sociedad de la información” [1] se define en relación a mecanismos como la producción, el tratamiento y la distribución de la información y, por ende, está directamente relacionada con el lenguaje. Esta sociedad exige, desde un punto de vista técnico, la infraestructura necesaria para la utilización de la información en

1 Puede consultarse la página del Gobierno de México: http://www.gob.mx/wb/egobierno/egob_conozca_la_concentracion_geografica_de_empres

todos los ámbitos de la economía y de la vida social. Es gracias a esta infraestructura que es posible encontrar información pertinente en el entorno diario que, en conjunto con la información en la memoria individual o colectiva, permite la construcción continua de conocimientos útiles para afrontar los retos cotidianos.

En la sociedad actual, que pone a nuestro alcance cantidades ingentes de información de todo tipo, una de las exigencias básicas a la que todo ciudadano se enfrenta es el aprovechamiento de esta información para construir conocimientos. Las actividades cotidianas involucradas en la construcción de conocimientos son las siguientes:

- Buscar información en diversas fuentes, tales como bibliotecas, diarios, enciclopedias, Internet, revistas, etc.
- Analizar la información que nos permita relacionarla, valorarla, seleccionarla y ordenarla en función de nuestras necesidades.

Estas actividades, reunidas bajo el concepto de “gestión de la información”, son indispensables para la construcción de nuevos conocimientos. Las sociedades de la información basan su valor en su capacidad para obtener, clasificar, sistematizar, sintetizar, difundir y aprovechar la información. Toda persona en esta nueva sociedad es productora de información y una adecuada gestión de la misma permitirá su aprovechamiento por cualquier otra persona.

Es importante destacar que, dada la complejidad de estas tareas, es prácticamente imposible realizarlas de forma manual; de ahí la necesidad de apoyarse en métodos automáticos.

El conocimiento ha estado presente en sistemas que operan de manera particular frente a cada problema; a diferencia de la información, la cual se emplea en sistemas rígidos cuya acción es general para los problemas que resuelven [2]. La operación central que se realiza en el conocimiento es la inferencia, esto es, se parte de ejemplos para generalizar el conocimiento.

Las formas en que se representa el conocimiento son variadas; principalmente relaciones, reglas, grafos, y jerarquías [3]. Por consecuencia, se requiere de especialistas, científicos computacionales

para llevar a cabo la realización e implementación de las diferentes representaciones de conocimiento en ambientes artificiales, es decir mediante sistemas de cómputo.

Un subconjunto de este tipo de sistemas, referidos como sistemas basados en conocimiento lingüístico (SBCL), es el formado por aquellos que utilizan conocimiento recopilado en léxicos, reglas morfosintácticas, heurísticas semánticas, etc. [4] La siguiente sección expone la importancia y beneficios de dichos sistemas, todo ello englobado en las llamadas tecnologías del lenguaje.

TECNOLOGÍAS DEL LENGUAJE

El lenguaje es el medio principal a través del cual los humanos intercambian información. Ningún otro sistema de información es tan eficaz para compartir ideas, transmitir conocimientos, expresar emociones y crear, en definitiva, lo que se conoce como cultura.

Las llamadas tecnologías del lenguaje constituyen la base de la ingeniería lingüística, entendida como la aplicación de los conocimientos sobre la lengua en el desarrollo de sistemas informáticos que puedan reconocer, comprender, interpretar y generar lenguaje humano en todas sus formas. Por lo anterior, el primer punto a destacar es que el desarrollo de las tecnologías del lenguaje requiere de equipos multidisciplinarios de investigación y desarrollo, en los que participan científicos en computación, inteligencia artificial, lingüistas, terminólogos e ingenieros, entre otros muchos especialistas.

En la actual sociedad del conocimiento, el uso de herramientas y servicios como portales de voz, correo electrónico leído a través del teléfono celular, dictado de documentos a la computadora, consulta de páginas en Internet en diferentes lenguas mediante traductores automáticos, e incluso tareas tan comunes como el uso de correctores automáticos y gramaticales durante la elaboración de documentos, se ha convertido en un conjunto de actividades cotidianas realizadas por millones de personas en todo el mundo. Sin embargo, no existe conocimiento mínimo de las necesidades económicas, técnicas y culturales requeridas para el desarrollo y mejora de estas herramientas, producto de la investigación y desarrollo en materia de tecnologías del lenguaje.

Las tecnologías del lenguaje son todas aquellas que se integran en aplicaciones informáticas para permitir el tratamiento de textos escritos —como el caso de la traducción automática o la búsqueda y recuperación de información—, o el procesamiento del habla —requerido para el dictado automático de documentos o la lectura de un correo electrónico en voz alta por la computadora—. Se trata, en conjunto, de tecnologías que hacen posible la creación de herramientas pensadas para auxiliarnos en el uso de las computadoras —y en general de las máquinas— sin renunciar por ello al uso habitual del lenguaje como medio de interacción y de intercambio de información [5].

Estas tecnologías abren un nuevo capítulo en la historia de la Sociedad de la Información; pueden enriquecer nuestro entorno informático con muchas de las funciones expresivas naturales propias del lenguaje humano, que hasta ahora no habían estado presentes en la interacción entre personas y máquinas. Combinando el potencial tecnológico con la profundidad y la flexibilidad del procesamiento del lenguaje humano, las aplicaciones y los servicios que ofrecen las tecnologías de la lengua pueden mejorar radicalmente la eficacia y la facilidad de los mecanismos y sistemas usados en las actividades cotidianas relacionadas con la información y la comunicación. Esto incluye el ocio y los negocios, la administración y la educación, los servicios y la vida diaria. Algunas aplicaciones típicas de las tecnologías del lenguaje en la actualidad son:

- Servicios de traducción automática, a los que se puede acceder a través de una página web o del teléfono móvil.
- Sistemas de dictado automático, que permiten transcribir cartas, informes o notas en diversos campos profesionales.
- Buscadores electrónicos con la integración de recursos lingüísticos, capaces de localizar y suministrar de forma inmediata información actualizada, a partir de bases documentales multilingües.
- Pantallas hablantes, que permiten a usuarios con discapacidades visuales superar las limitaciones que habían tenido hasta ahora.
- Diccionarios electrónicos multilingües, que facilitan la lectura de textos en lenguas extranjeras.

Como se ha mencionado, las tecnologías del lenguaje involucran una serie de técnicas relacionadas con el tratamiento del lenguaje por medios computacionales. Pueden distinguirse dos grandes vertientes, las que se aplican al tratamiento del lenguaje escrito y aquellas para el tratamiento del lenguaje oral o hablado. Las siguientes secciones profundizan en dichas vertientes [6].

TECNOLOGÍAS PARA TRATAMIENTO DE TEXTO

Actualmente, el acceso a la información debe ser selectivo y guiado por objetivos; si no es así, a pesar de disponer de una gran cantidad de datos al alcance, no se llegarán a convertir en información que se pueda integrar como conocimiento. Este hecho plantea la necesidad de definir métodos de almacenaje y codificación de la información a fin de que posteriormente pueda ser debidamente recuperada.

El objetivo último es que los documentos puedan ser indexados por conceptos, que se pueda acceder a la información mediante interacciones en lenguaje natural y que sea posible interactuar y recuperar la información en la lengua deseada, todo esto independientemente de la lengua en la que se encuentre originalmente la fuente de consulta.

Sin embargo, actualmente los documentos se indexan por formas o por palabras clave, sin tener en cuenta las relaciones entre las variantes flexivas de las palabras o las relaciones de significado. Tanto las empresas editoriales, en cuanto creadoras de contenidos estructurados (diccionarios y enciclopedias), como los sistemas de búsqueda en Internet o la gestión de los grandes bancos de datos requieren sistemas que permitan el acceso al contenido desde diferentes perspectivas, según las necesidades de los usuarios.

Entre otras técnicas asociadas a la gestión de información se encuentran la extracción de información, la recuperación de información, la traducción automática y la elaboración de resúmenes. Todas ellas requieren una serie de recursos para el procesamiento de los datos, entre los que destacan los analizadores morfológicos, los desambiguadores, los analizadores sintácticos, las redes semánticas, los corpus etiquetados y los lexicones computacionales y multilingües.

El desarrollo de los distintos sistemas que utilizan de los recursos de ingeniería lingüística permite optimizar los resultados de las búsquedas. Las mejoras se pueden producir en diferentes sentidos:

- Permitiendo expresar las búsquedas en lenguaje natural y sin restringirlas a un metalenguaje, que es una combinación de palabras clave y operadores booleanos de aprendizaje complicado.
- Ampliando el espacio de búsqueda mediante la ampliación de la pregunta del usuario incorporando en él sus formas flexivas y las palabras semánticamente relacionadas.
- Restringiendo los documentos seleccionados a aquellos que realmente responden al sentido de las búsquedas.

TECNOLOGÍAS PARA EL TRATAMIENTO DEL HABLA

El uso de la lengua oral como modo de comunicación con los sistemas informáticos permite, en primer lugar, liberar las manos y la vista, lo que facilita la recepción de la información desde cualquier punto y hace posible la movilidad, de manera que pueden realizarse otras actividades simultáneas. La entrada de datos en la gestión de los fondos de un almacén, las operaciones relacionadas con el mantenimiento y el control de la calidad en entornos industriales, la recepción de información en entornos como cabinas de avión —con una gran cantidad de sistemas de control e indicadores—, el control de la radio, del teléfono o de la climatización de un coche y la obtención de información sobre itinerarios —el prototipo desarrollado en el proyecto *Vodis*² (*Advanced Speech Technologies for Voice-operated Driver Information Systems*) ofrece estas posibilidades— o bien el dictado de documentos son algunas de las aplicaciones para las cuales se ha demostrado que la utilización de la lengua hablada es una ventaja importante.

En segundo lugar, utilizar el habla para acceder a la información almacenada en un sistema de cómputo brinda la posibilidad de usar el teléfono en vez de la pantalla y el teclado tradicionales, con todo lo

2 <http://cordis.europa.eu/ist/98vienna/xvodis.htm>

que ello implica en cuanto a rapidez. Hay que pensar, además, en la independencia que este procedimiento proporciona con respecto a la distancia y la disponibilidad de una computadora. Algunas aplicaciones en este ámbito son los servicios de información sobre horarios de transportes públicos, las consultas en las páginas amarillas u otros directorios, la información ciudadana, la banca telefónica o las centrales automatizadas.

Finalmente, no hay que olvidar que guardar y recibir información mediante el habla es una alternativa práctica cuando hay que prescindir, debido al tamaño, del teclado, como sucede con las agendas electrónicas y los comunicadores personales. La voz, como los lápices electrónicos y las pantallas táctiles, es una modalidad viable para hacer aumentar la funcionalidad de estos aparatos. Además que esto es un nicho de mercado abierto para aplicaciones en la lengua española.

Hoy en día disponemos de programas comerciales que posibilitan el dictado automático, por ejemplo los desarrollados por Dragon Systems, IBM o Philips, la navegación por Internet, por ejemplo el Free Speech Browser³, y la recuperación de información mediante el habla. También es posible la lectura en voz alta de un texto en formato electrónico, como podría ser un mensaje de correo o el contenido de una base de datos. Es posible, además, con los sistemas de diálogo que incorporan síntesis, reconocimiento y un cierto grado de comprensión, obtener información o efectuar transacciones comerciales mediante el teléfono.

Estos últimos son, seguramente, sistemas que en la actualidad despiertan más interés y, por este motivo, se mencionan algunas de las aplicaciones desarrolladas en el contexto de la comunidad europea, donde el fenómeno del acceso a la información multilingüe es crucial para el funcionamiento adecuado y fluido de gran parte de la vida cotidiana.

Entre otras aplicaciones de las tecnologías del habla podemos mencionar los sistemas de diálogo (algunos complementados con información multimodal), la información ciudadana de forma automática mediante el teléfono, los sistemas de traducción automática de con-

3 http://findarticles.com/p/articles/mi_m0NEW/is_1999_Nov_4/ai_57296080

versación, tanto telefónica como cara a cara, la enseñanza de lenguas asistida por ordenador y la ayuda para los sectores con necesidades especiales.

Tal como queda manifiesto, al incorporarse al mundo de la informática y las telecomunicaciones, las tecnologías del lenguaje están creando un nuevo mercado para productos y servicios especializados.

IMPORTANCIA DE LAS TECNOLOGÍAS DEL LENGUAJE EN EL MANEJO DE INFORMACIÓN

La clave para el manejo futuro de la información y el conocimiento son los textos. Una prueba de ello es el impacto que tiene Internet en el trabajo cotidiano a todos los niveles, pues gestiona el acceso al conocimiento aunque todavía de manera rudimentaria. En este sentido, la gestión de textos resulta prioritaria en el actual desarrollo de las tecnologías de la información e incide favorablemente en el crecimiento tecnológico de la nación.

El tratamiento de información oral y escrita requiere de estudios teóricos del lenguaje natural (realizado por profesionales como son lingüistas, bibliotecólogos, comunicólogos y otros pertenecientes al área humanística) así como de una serie de metodologías y técnicas de análisis (lo que involucra a profesionales de campos como computación, informática, estadística, ingeniería, etc., con un perfil de ciencias exactas). En este sentido, la interdisciplina incide favorablemente en la interacción y colaboración de sectores del conocimiento que, desde un punto de vista tradicional, se han visto como desligados [7]. Por ello, resulta vital ayudar a aumentar el crecimiento de una sólida planta de investigadores, docentes y técnicos competitivos tanto para el sector de las ciencias exactas como el de las ciencias humanísticas.

El fruto de esta fusión de disciplinas ya se ha visto concretado en el importante auge que ha tenido en otros países lo que se conoce como *tecnologías e industrias del lenguaje y la información*. Este sector ha dado lugar a una gran cantidad de áreas de investigación y estudio fundadas desde mediados del siglo pasado. Estas áreas se han consolidado hasta lograr un crecimiento exponencial en los 80 y 90. De entre ellas, figuran entre las más relevantes la inteligencia artificial, la robó-

tica, el aprendizaje-automático, el procesamiento de lenguaje natural, la ingeniería lingüística, la lingüística computacional, la extracción y recuperación de información, el procesamiento de voz, la representación y generación de conocimiento, la terminótica y la minería de textos. En el caso de nuestro país, la relevancia estratégica que tienen todas estas áreas es aún emergente, debido en gran medida, a la falta de conocimiento y difusión. Por ello, para lograr un alto grado de calidad y competitividad en la generación y aplicación de tecnologías de la información en el país, es prioritario fomentar y destinar fondos para el desarrollo de estas interdisciplinas en nuestro medio.

La creación de una plataforma para el procesamiento del lenguaje natural en español, así como el desarrollo de recursos lingüísticos en nuestro idioma, como son los corpus lingüísticos, los bancos terminológicos, los tesauros y las ontologías, será de alto provecho para la generación propia de múltiples aplicaciones como sistemas de resumen automático, interfaces inteligentes, sistemas pregunta-respuesta, sistemas de extracción de información, buscadores inteligentes, entre otros de los que actualmente se empiezan a crear en el país.

Los beneficios de contar con tecnologías del lenguaje y el aprovechamiento de sus recursos derivados son varios para la gestión de información, entre los que podemos encontrar:

Búsqueda inteligente de información, que permita a los usuarios de bases de datos documentales realizar búsquedas textuales en dichos repositorios de información, utilizando frases y contexto, reduciendo con ello las deficiencias de los buscadores actuales, las múltiples búsquedas realizadas manualmente y el gran número de resultados carentes de sentido a los requerimientos de información.

Distribución selectiva de información a través de asistentes automáticos personales encargados de monitorear fuentes de información en línea, detectar nueva información, enviar una notificación al usuario con un resumen de dicha información y los accesos necesarios para que pueda consultarla.

Poblado automático de bases de datos, que permite a los usuarios obtener, a partir de textos (por ejemplo, noticias, informes, etc.), conjuntos de datos de forma automática, como nombres de personas y sus cargos, nombres de eventos, cifras relevantes, etc. Estos datos

son almacenados en bases de datos para la generación de estadísticas, análisis de datos o gráficas.

SITUACIÓN ACTUAL EN MÉXICO

La falta de un proyecto a nivel nacional que rijan, gestione y coordine el trabajo de Investigación y Desarrollo (I+D) de los grupos, laboratorios y centros relacionados con las tecnologías de la información provoca una enorme desorganización e inhibe avances en la solución de problemas específicos. La consecuencia lógica es un estancamiento tecnológico y un elevado gasto para el país, pues se ve obligado a adquirir tecnología extranjera ante la poca o nula producción nacional, esto sin mencionar los altos costos de implementación y difusión de tecnologías de información que de inicio son ajenas a la cultura e idiosincrasia de los usuarios del país. Un ejemplo de esto son los llamados sistemas para inteligencia de negocios, diseñados y producidos principalmente por firmas de tecnologías de información extranjeras, que combinan complejos mecanismos computacionales para la explotación de los datos almacenados en sistemas empresariales tradicionales u otras fuentes potenciales de información; sistemas que además requieren de inversiones importantes en el orden de varios cientos de miles de pesos para su adquisición e implantación en territorio nacional. Este tipo de herramientas son inaccesibles y en el mayor de los casos prohibitivas para el grupo empresarial más representativo en México, las micro, pequeñas y medianas empresas [8].

Otro factor que debe ser abordado es la gran deficiencia formativa y de visión interdisciplinaria a nivel académico, aunado a la carencia en la promoción laboral para profesionistas egresados tanto del nivel medio-superior (técnico), superior (licenciatura) y postgrado (maestría, doctorados y post-doctorado). Cabe señalar que a nivel nacional, un gran porcentaje de las carreras afines a las Tecnologías del Lenguaje, la Información y el Conocimiento (TLIC) carecen de un perfil formal y vinculado a la industria para la investigación aplicada, el diseño y desarrollo de nuevas herramientas de producción nacional y en beneficio directo del consumo y los mercados internos.

Aunado a los factores expuestos, el papel que juega la sociedad mexicana sigue siendo uno de los temas decisivos en el crecimiento

y desarrollo de las TLIC en México. La sociedad, siendo primordialmente consumidora, está habituada a la adquisición de tecnología aún cuando ésta no haya sido diseñada para resolver sus necesidades particulares. Como consecuencia, la demanda por nuevos y mejores productos nacionales, así como el uso de las TLIC se ve obstaculizada por el hecho de que los usuarios de éstas deben adecuarse a la tecnología y en muy pocas ocasiones la tecnología se adecua a las necesidades de los usuarios o de los potenciales consumidores. Otro perfil de este problema es que la sociedad también se ha habituado y considera, en un porcentaje muy amplio, que los productos y servicios de las TLIC no merecen una remuneración económica, desvirtuando la industria y provocando que, ante el fenómeno de la piratería, sólo las grandes empresas transnacionales gocen de algún factor de crecimiento, mientras las pocas empresas nacionales sobreviven -en el mejor de los casos- del mercado nacional.

Finalmente, es importante reconocer y tratar el problema relativo a los recursos humanos ubicados en las organizaciones que representan los usuarios potenciales de las TLIC. Al tratarse de tecnologías emergentes, estos recursos humanos no están preparados para aceptar dichas tecnologías, ya sea por desconocimiento de las mismas, o bien por la falta de habilidades y capacitación para llevar a cabo las implementaciones necesarias. Como consecuencia, se opta por la adquisición de otro tipo de tecnologías que tal vez no resuelven el problema de los usuarios, o que sólo ofrecen soluciones parciales, además de implicar costos más elevados ya sea en el corto, mediano o largo plazo, dado que generalmente dichas soluciones provienen de empresas extranjeras.

EL FUTURO EN MÉXICO

Dada la situación actual descrita, resulta necesario concretar un proyecto a nivel nacional para contribuir a la gestión de información que conduzca al descubrimiento, generación y administración de conocimiento, con base en el uso de tecnologías de lenguaje, partiendo de la inmensa base de información existente en medios electrónicos tanto en formatos textuales como orales.

Una propuesta para ello es crear un centro de ámbito nacional para la investigación aplicada y desarrollo tecnológico, con carácter formativo y de vinculación empresarial, que encamine los esfuerzos para la creación de nuevas tecnologías para la gestión de información y conocimiento con base en las tecnologías del lenguaje. El fin último será coordinar los diversos esfuerzos que existen en I+D en el área para que converjan hacia un fin común, mediante un plan de desarrollo estratégico conforme a las necesidades del país. Entre los objetivos particulares de esta propuesta se proponen:

- Incrementar la I+D nacional de productos para la gestión y explotación de información científica, tecnológica, administrativa y legal para llevar a cabo los más diversos objetivos y, especialmente, para generar conocimiento a fin de alcanzar niveles competitivos y de calidad, logrando con ello un impacto a nivel internacional.
- Facilitar los medios para captar y usar de un modo inteligente la información oral, escrita o visual, proveniente de cualquier campo del conocimiento y de cualquier sector, sea público, privado o social.
- Encauzar la investigación básica y aplicada para crear tecnologías capaces de procesar y manejar volúmenes gigantescos de información que se requieran en la generación de conocimiento.
- Desarrollar productos y recursos tecnológicos (por ejemplo, corpus lingüísticos, bancos terminológicos, ontologías y tesauros) que faciliten el procesamiento de información requerida para fines específicos.
- Poner al alcance de las empresas e instituciones, así como de la población en general, sistemas para recuperar y extraer información, mediante el uso de lenguaje natural.
- Proporcionar los canales adecuados para el consumo de tecnología desarrollada dentro del país a través de una adecuada gestión de cambio organizacional que dé apoyo a este tipo de propuestas.

- Vincular los sectores públicos, privados y sociales para el fomento del desarrollo y uso de las tecnologías, recursos y productos desarrollados en el país. Asimismo, coadyuvar al acercamiento entre el ámbito académico y el ámbito empresarial para la emergencia de colaboraciones entre los dos rubros y para la optimización de la tecnología desarrollada.
- Fomentar la formación de profesionales competitivos y de recursos humanos de alta especialización en el país, sobre todo considerando que pueden llegar a ser más competitivos y rentables que sus similares de países del primer mundo. En este rubro se crearán programas de licenciatura y posgrado que apoyen la especialización y la competitividad de los integrantes del proyecto y de los cuadros que a largo plazo trabajen en este rubro para el gobierno y las empresas.
- Generar ingresos mediante la venta en el extranjero de nuestra tecnología, en lugar de causar gastos en la compra de la misma.
- Cabe señalar que la creación de un centro de esta naturaleza fue propuesta recientemente al Consejo Nacional de Ciencia y Tecnología por iniciativa del Instituto de Ingeniería de la UNAM, con la participación de grupos de investigación de diversas instituciones en el país, como son el Instituto Politécnico Nacional, el Instituto Nacional de Astrofísica, Óptica y Electrónica, la Benemérita Universidad Autónoma de Puebla, El Colegio de México, y de la misma UNAM.

REFERENCIAS

- [1] R. Trejo Delarbre, "Vivir en la Sociedad de la Información. Orden global y dimensiones locales en el universo digital," en *Comunicación Educativa en la Sociedad de la Información*, (Ed. R. Aparici), Madrid: Universidad Nacional de Educación a Distancia de España, 2003.

- [2] Norvig, Peter y Russell, Stuart, *Artificial intelligence: a modern approach*, (2nd ed.) , Prentice-Hall International, New Jersey, 2003.
- [3] John F. Sowa: Knowledge Representation: *Logical, Philosophical, and Computational Foundations*. Brooks/Cole: New York, 2000.
- [4] Mitkov, Ruslan, *The Oxford Handbook of Computational Linguistics*, Oxford University Press, New York, 2003.
- [5] Cole R.A., J. Mariani, H. Uszkoreit, G. Varile, A. Zaenen, V. Zue, A. Zampolli. *Survey of the State of the Art in Human Language Technology*, Cambridge University Press and Giardini, Cambridge, 1997.
- [6] M. A. Martí y Joaquim Llisterri, “La ingeniería lingüística en la sociedad de la información”, *Digitum, Revista digital d’humanitats* (3), 2001.
- [7] G. Sierra y A. Medina, “Ingeniería Lingüística: Interdisciplina aplicada al desarrollo humano” en *Memorias del Primer Congreso de la Academia de Ingeniería*, Academia de Ingeniería, México, 2003.
- [8] L. Orozco, “Sistemas y tecnologías de información en las micro y pequeñas empresas”, *Razón y Palabra*, núm 39, junio-julio 2004.