

Importancia de la terminología para localizar y recuperar información científica.

IMELDA VELÁZQUEZ MONTES

Facultad de Química.

Universidad Nacional Autónoma de México.

INTRODUCCIÓN

La información digital actualizada es un medio muy importante, para que los investigadores puedan mantener sus trabajos en la frontera del conocimiento; en virtud de la velocidad con la que se publican y difunden los descubrimientos en el ámbito científico, la importancia en el manejo de esta información y la rapidez con la cual se puede tener acceso a ella, es de gran beneficio para los investigadores, profesores y estudiantes; su uso permite incrementar su productividad, colaborando de ese modo al incremento del conocimiento.

En México es posible lograr estos beneficios, en la Universidad Nacional Autónoma de México y específicamente en la Facultad de Química, dentro de la Hemeroteca se ofrecen Servicios de Información Digital (SID), que tiene como objetivo principal ampliar y diversificar los servicios de información digital principalmente a la comunidad universitaria.

En estos servicios, no solo se encuentra la información de divulgación nacional, sino también internacional ya que cuenta con acceso a bases de datos a nivel mundial. Donde el principal objetivo es poder brindar un buen servicio, donde se obtenga de manera ágil y rápida

información actualizada en la mayor cantidad de bases de datos, para un mejor aprovechamiento de las mismas.

La información química que se produce en el mundo es enorme, se calcula en más de las tres cuartas partes del total de la información científica en todas sus modalidades; lo que equivale a montañas de documentos de Química o de cualquiera de las áreas relacionadas con la misma. Sería imposible que una persona pudiera revisar la información, ni siquiera trabajando las 24 horas del día los trescientos sesenta y cinco días del año. Es por eso que las bases de datos que tratan de incluir toda esta información, ya no pueden ser publicadas en formato impreso; un ejemplo es la base de datos más grande a nivel mundial, llamada Chemical Abstracts, que próximamente dejará de imprimirse, ya que por su gran tamaño, no es posible seguirse consultando en forma impresa. Desde hace tiempo el Chemical Abstracts se produce en varios formatos, como son los discos compactos, la versión llamada STN Easy, etcétera, hasta llegar a la producción de softwares especiales para el manejo de esta información, como el programa llamado STNEXPRESS y recientemente se diseñó un software más amigable llamado SciFinder Scholar, diseñado para que el usuario final lo utilice de forma fácil y dinámica.

A pesar de la facilidad que puede brindar el uso de herramientas tecnológicas para la localización y recuperación de información científica tan diversa y abundante, se presentan los siguientes problemas:

- 1 Si el investigador que publica los resultados de sus trabajos no utiliza la terminología apropiada, que los editores llaman "Key words" o palabras clave, para que sirvan como un verdaderos descriptores del conocimiento producido, de manera que permita ser identificado plenamente, para ser buscado y consultado por otros investigadores o lectores, su trabajo corre el riesgo de quedar en el olvido, simplemente porque nadie más podrá localizarlo.
- 2 Si los editores no establecen criterios claros y precisos de indexación y clasificación de los artículos publicados, sobre todo en formato digital, de tal forma que se les pueda localizar con acierto, usando para ello la terminología apropiada, este hecho también

puede provocar que los trabajos científicos sean enterrados para siempre.

3 Las llamadas nuevas Tecnologías de Información y Comunicación (TIC) proporcionan herramientas útiles para que los usuarios con cierta facilidad busquen y encuentren en las bases de datos, los artículos que necesitan para realizar sus investigaciones, pero, que pasa al no utilizar la terminología adecuada para ello, se corre el riesgo de no poder acceder a documentos que puedan ser clave para la realización de su trabajo.

4 En el caso de la Química, donde las reglas de nomenclatura son complicadas y difíciles de entender, porque su organización lógica se basa en las propiedades de los compuestos, esto es, en el forma de reaccionar dependiendo de las condiciones en las que se encuentran; por lo tanto, la localización de un compuesto se dificulta demasiado, sobre todo al no utilizar la terminología precisa.

5 Considerando que al pasar el tiempo, existe cada día mayor interacción entre las ciencias y cada día se van perdiendo las barreras entre ellas, de manera que ahora se sintetizan compuestos órgano metálicos que tienen actividad biológica o fisiológica, incidiendo en la vida social y económica de los países; esto genera una situación complicada para el buscador de información, ya que aún cuando se usen bases de datos multidisciplinarias, actualmente no existe la base de datos que agrupe varias especialidades, como Química, Biología, Medicina, economía y sociología; por tanto, si no se usan las bases de datos y los términos adecuados, la consulta fracasará.

La finalidad de este trabajo es plantear las diferentes formas con las que se pretende abordar la problemática existente, de manera que no sea un obstáculo para localizar y recuperar información requerida, en todas sus modalidades.

Partiendo del hecho de que las ciencias, en su mayoría cuentan con un lenguaje especializado, inherente a las mismas, además de que el conocimiento está sistemáticamente organizado en cada una de ellas.

En el caso de la Química existen verdaderos tratados, que intentan organizar en forma lógica y característica a todos los compuestos

existentes; se plantea como una alternativa, el hecho de que los buscadores de información deben ser expertos en la ciencia donde pretendan localizar y recuperar información, así en el caso de la Química se observa que se contratan personas con posgrado en Química, con la particularidad en el manejo de la terminología apropiada y de los sistemas de organización de la información, para especializarse como buscadores de dicha información.

DESARROLLO

Los servicios de información digital (SID), proporcionan servicios de información competitivos a nivel nacional e internacional que apoyan los programas académicos y de investigación, promoviendo los valores y habilidades de los estudiantes y profesionales, definidas en la misión de la Universidad.

Basándose en catálogos electrónicos y bases de datos que integran los servicios de información bibliográfica digital que adquiere la Dirección General de Bibliotecas (DGB), se ofrecen los nombres de las diferentes bases de datos a la comunidad universitaria, presentándolas en un formato accesible para el usuario.

Algunos sistemas contienen artículos de texto completo y otros se encuentran relacionados con algún servicio de documentación que permite la obtención inmediata del documento localizado en la base de datos. La información que se puede obtener en este servidor es muy variada, ya que cuenta con información de diferentes bases de datos, revistas electrónicas y catálogos globales; ayudándose del uso de nuevas tecnologías para poder realzar una búsqueda más ágil y rápida.

Es indispensable que la(s) persona(s) encargada(s) de brindar los servicios propuestos, tengan la suficiente preparación, tanto en el manejo de la terminología química, como en los sistemas de nomenclatura de los compuestos sintetizados en toda la historia de la humanidad, para tener éxito en dicha empresa.

RESULTADOS

En los catálogos de las bases de datos se encuentra contemplada la siguiente información:

- Nombres de las bases de datos: Se presentan en la lengua y con las características tipográficas originales.
- Cobertura: Se refiere a los años que cubre la información contenida en cada base de datos.
- Número de registros: Proporciona el tamaño de la base de datos, ayudando al usuario a saber escoger entre los datos que aparecen y la información que necesita.
- Sistema: Ubica los conjuntos de bases de datos que pueden ser consultados con un mismo software de recuperación de información.

Entre todos ellos están el autor del artículo, él o los temas, la fecha de publicación, desde luego, el título de la revista, el título de los artículos y otros.

Como por ejemplo:

Base de datos	Cobertura	No de revistas o registros incluidas	Sistema en que se encuentra	Texto completo
34.Health Source Plus Esta base de datos contiene información especializada en temas de nutrición, ejercicio y cuidados médicos, fármacos y alcoholismo.	1984 a la fecha	Más de 200 revistas 400 000 artículos	EBSCO	Si

Dentro de las principales bases de datos se encuentra “Elsevier” publicada por uno de los más importantes editores científicos del mundo, así como Chemical Abstracts, siendo esta última para el área científica, la base de datos más completa, que tiene registros de más

de once mil títulos de revistas internacionales arbitradas, patentes y libros, ésta base, se constituye en la base de datos más consultada tanto para localizar información como para encontrar la totalidad de las citas de trabajos publicados en el ámbito de las ciencias exactas.

Otro ejemplo:

Base de datos	Cobertura	No de revistas o registros incluidas	Sistema en que se encuentra	Texto completo
22.Dissertation Abstracts Contiene todas las tesis doctorales presentadas en los últimos 137 años en instituciones reconocidas en los Estados Unidos; Contiene tesis de maestría y tesis de doctorado en idiomas extranjeros	1861 a la fecha	1400000 registros	OCLC	Una parte es gratis. Se puede comprar la tesis completa

TÍTULOS DE REVISTAS ELECTRÓNICAS DE CONSULTA DIRECTA

Los creadores y los editores de recursos de información en red, no utilizan ni formatos, ni visualizadores iguales, tampoco organizan los datos de manera normalizada. Es por ello que, los servicios que incluyen servicios digitales varía considerablemente; mientras que algunos permiten el acceso directo a un título, y la búsqueda de los artículos en cada fascículo, semejando a los formatos impresos, en otros casos la información se organiza en base de datos con múltiples accesos en páginas web que tiene poco parecido con la revista en papel.

Por lo anterior, la Dirección General de Bibliotecas (DGB), agrupa mas de quince mil títulos de revistas electrónicas en un catálogo que facilita el acceso directo a las mismas. En dicho catalogo el lector puede buscar un titulo, un tema, un editor o una palabra del titulo, cuando no recuerda la palabra inicial del mismo.

Entre las revistas que se encuentran en el proceso de instalación están los títulos de EBSCO, uno de los más grandes editores del mundo.

Ejemplo de revistas electrónicas de consulta directa:

- ABSTRACTS-PERIODICAL OF THE AMERICAN PHYSIOLOGICAL SOCIETY
- AMERICAN JOURNAL OF PHYSIOLOGY, CELL PHYSIOLOGY
- BIOCHEMICAL AND MOLECULAR MEDICINE
- JOURNAL OF INORGANIC BIOCHEMISTRY
- JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
- SCIENCE
- NUCLEAR PHYSICS B
- WEED RESEARCH
- WATER SCIENCE AND TECHNOLOGY

TÍTULOS DE REVISTAS ELECTRÓNICAS DE CONSULTA INDIRECTA

La información que se puede obtener en este tipo de servicios es muy variada; entre los puntos de acceso se encuentra el autor del artículo, el o los temas, la fecha de publicación, desde luego el título de la revista, el título del artículo y otros.

Al recuperar información pertinente en la base de datos, el lector podrá encontrar la opción de visualizar el texto completo e imprimirlo, grabarlo o enviarlo a un correo electrónico.

A esta información se le conoce como consulta indirecta, porque no presenta los datos de manera semejante a la que contiene la revista impresa.

Ejemplo de revistas electrónicas de consulta indirecta:

ABA JOURNAL
ADVOCATE
AMERICA'S NETWORK
HUMANIST
HECATE
MANUFACTURING CHEMIST
MEDICAL TEACHER
SOCIOLOGY
STUDENT BMJ

Los nombres y direcciones de las páginas llegan continuamente a manos de la Profesora Imelda Velázquez Montes por medio de ligas

mundiales en Internet o comunidades virtuales que se dedican a la divulgación de información científica, donde a través de un servidor se manda la información más actualizada acerca de los temas en los cuales está interesado el usuario, que en todos los casos es información científica, en especial relacionada con la química; con estos datos, se verifica y seleccionan los sitios que contienen información útil y específica, ya que al hacer la selección se valúa el contenido de éstas para saber en qué temas se relacionan para así poder colocarlos en los sitios específicos, para que el usuario los pueda buscar posteriormente.

Es por eso que el trabajo que realiza el departamento de Servicios de Información Digital (SID), ayuda de manera titánica en la búsqueda y recopilación de información, proporcionando los medios para facilitar la obtención de información actualizada y veraz, a todo el usuario que lo solicite.

A partir del presente año SciFinder Scholar 2007 ofrece nuevas y poderosas herramientas:

- **Guardar Respuestas**— Ahora usted puede guardar listas de referencias, sustancias, o reacciones y re-utilizarlos mas tarde en otra sesión de SciFinder Scholar.
- **Combinar conjuntos de Respuestas**—La herramienta de combinar hace posible enfocarse en un tópico de investigación particular de una nueva forma, combinando sustancias, reacciones y referencias en un conjunto de respuestas activo con aquellos que hayan sido guardados previamente. Empleando varias opciones de combinación, intersección, o remoción, usted puede manipular y personalizar la información contenida en esos conjuntos combinados de respuestas.
- **Exportar fichas comerciales desde CHEMCATS® en Excel**— Crear reportes listando sus fuentes comerciales favoritas, incluyendo información de contactos y precios para sustancias de interés. La información del Catálogo se importa a hojas para su clasificación, manipulación y referenciación.
- **Explorar desde Desplegado de sustancias**—Llegue a ser más productivo usando el desplegado de una sustancia, para iniciar una nueva búsqueda o refinar una búsqueda activa.

- **Imprima Estructuras de sustancias en formato de tablas**—Obtenga una vista comparativa única de múltiples estructuras químicas, imprimiéndolas una junto a otra en este útil arreglo en forma de tabla. Compatible con Windows Vista™ and Mac OS X
- **Categorizar Referencias**—Esta herramienta le permite ordenar y evaluar información en su conjunto de respuestas, creando un histograma de fácil análisis. Categorizar separa referencias relacionadas por sus términos asociados por el Índice de CA. Con Categorizar usted puede re-armar un subconjunto de una forma única creando oportunidades para nuevas perspectivas

DISCUSIÓN

En cuanto a la terminología en el ámbito químico, al utilizar el mencionado Sistema SciFinder Scholar, se pueden usar las herramientas descritas durante la búsqueda de un compuesto específico; incluyendo el sistema de nomenclatura aprobado a nivel mundial, que incluye un número de registro llamado CAS Registry Number para cada sustancia, elemento o compuesto, que además de confiable puede eliminar equivocaciones o errores que se pueden cometer al utilizar la nomenclatura o un idioma diferente al del origen del documento, así como otro tipo de términos científicos incluidos en los documentos y que son clave para su identificación.

La razón por la cual es indispensable que la búsqueda y selección de información la realice un profesional de la química, de preferencia con posgrado es porque, cuando no se cuenta con el número CAS, ni la fórmula exacta, ni la estructura del compuesto; se hace necesario conocer las propiedades, las características químicas o bien su comportamiento en las condiciones de reacción, de manera que se garantice el éxito en la consulta a la base de datos adecuada.

Para el caso específico de localización del tema pigmentos o colorantes, fue posible hacer el siguiente comparativo:

El criterio de selección e los vínculos en buscadores y meta buscadores fue en función del contenido de la página vinculada; éste tenía que poseer al menos una de las siguientes características:

- Información de interés técnico
- Información de interés divulgativo
- Información de interés científico
- Información respecto al contexto cultural de los pigmentos

Excluyendo:

- Publicidad
- Textos en idiomas diferentes al español y al inglés
- Información de pigmentos de interés o aplicación estrictamente en procesos biológicos.

La información disponible en Chemical Abstracts fue seleccionada por su contenido:

- Información de interés técnico
- Información de interés científico

Excluyendo:

- Información en idiomas diferentes al español y al inglés
- Información de pigmentos de interés o aplicación estrictamente en procesos biológicos

Todos los artículos o patentes encontrados a través del Chemical Abstracts fueron de interés científico, tecnológico o ambos; se referían a tecnologías o investigaciones de punta al respecto del tema de pigmentos.

De estos datos, se puede concluir que con fines de recavar información divulgativa e introductoria para los temas se vuelve conveniente la búsqueda en fuentes de acceso libre, siendo la más efectiva Copernic; que si bien no arrojó muchos resultados, si brindó información de mejor calidad que Google.

Mientras que la búsqueda a través del Chemical Abstracts brinda información mucho más específica, actual e interesante al respecto de temas científicos y tecnológicos. Sin embargo, la única desventaja

que se puede encontrar en el Chemical Abstracts es que no fue posible recavar toda la información de interés en texto completo que se encontró debido a que está escrita en idiomas como el ruso, japonés, chino y coreano. Lo que me hace pensar que los adelantos realizados en el área de pigmentos en el mundo fueron principalmente desarrollados en países asiáticos; sin embargo, se localizaron todos los resúmenes o abstracts de los trabajos publicados durante 100 años, utilizando la terminología apropiada.

BIBLIOGRAFÍA:

Servicios de información digital para la comunidad universitaria. 1999 Universidad Nacional Autónoma De México. Dirección General de Bibliotecas. <http://www.unam.mx>

Imelda Velázquez-Montes, "Internet para todo y para todos" Antología 1, Ciclo de Conferencias La Ciencia más allá del Aula. FQUNAM. 137-150, (2006). 3. Portal de Internet [en línea] [lugar de publicación México, D: F.] [31/05/06 14:20 PM]. Disponible en World Wide Web: (www.mtas.es/ins-ht/ntp_276.htm)

Answer extraction and ranking strategies for definitional question answering using linguistic features and definition terminology. By: Han, Kyoung-Soo; Song, Young-In; Kim, Sang-Bum; Rim, Hae-Chang. *Information Processing & Management*, Mar2007, Vol. 43 Issue 2, p353-364, 12p; DOI: 10.1016/j.ipm.2006.07.010; (AN 22946071)

Archival Terminology; or, What Did You Just Call Me? By: Sandleback, Nancy. *MAC Newsletter*, Jan2007, Vol. 34 Issue 3, p39-39, 2/3p; (AN 23799323)

A Glossary of Archival and Records Terminology. By: Robertson, Anne. *Archives & Manuscripts*, Nov2006, Vol. 34 Issue 2, p203-204, 2p; (AN 23583760)