

Estudios de solapamiento en la selección de las publicaciones seriadas y las bases de datos

Rubén Cañedo Andalia¹

Resumen

La selección de publicaciones seriadas y bases de datos resulta una tarea sumamente difícil, en especial, para el sector de la biomedicina, por el crecimiento acelerado de las fuentes y los medios de información en esta área del conocimiento. En tal sentido, la realización de investigaciones informétricas es una tendencia actual. Dentro de ellas, los estudios de solapamiento representan una forma útil para determinar el grado de procesamiento que reciben las publicaciones de una temática específica en distintas bases de datos, a la vez que permiten hallar el nivel de exhaustividad que éstas presentan en el cubrimiento de la literatura de una rama del saber. Se exponen los fundamentos teórico-metodológicos de los estudios de solapamiento, algunas formas prácticas para su aplicación tanto en la evaluación de los títulos de publicaciones seriadas como de las bases de datos, así como sus principales ventajas y limitantes.

Descriptor: PUBLICACIONES SERIADAS; BASES DE DATOS; SELECCIÓN; SOLAPAMIENTO; INFORMETRÍA/tendencias.

Los análisis de solapamiento permiten determinar el grado de semejanza que presentan 2 o más bases de datos en el procesamiento de artículos, revistas, instituciones, países y otras entidades. Dichos análisis resultan de gran utilidad, tanto para la selección de bases de datos como para la obtención de publicaciones seriadas de gran procesamiento, divulgación y utilización.

Se considera una publicación solapada cuando ésta es cubierta al menos en un producto impreso o automatizado de 2 o más servicios de información secundaria.

Las revistas que presentan algún grado de solapamiento constituyen generalmente un recurso esencial para el trabajo de las comunidades científicas que laboran en la corriente principal de desarrollo de una o varias materias. Dichos títulos poseen, a su vez, un mayor efecto multiplicador.

Una referencia bibliográfica o una publicación posee efecto multiplicador cuando se procesan por 2 o más servicios de información secundaria, cuando satisfacen necesidades planteadas a 2 o más servicios

¹ Licenciado en Información Científico-Técnica y Bibliotecología. Departamento Recursos de Información. Centro Nacional de Información de Ciencias Médicas.

distintos en una misma institución de información o, simplemente, cuando responden a un grupo de perfiles diversos y de interés a una comunidad de usuarios determinada. Sin embargo, a los efectos del estudio que se presenta, se aplican las definiciones de solapamiento y efecto multiplicador sólo a aquellas referencias y títulos que procesan 2 o más productos (bases de datos o servicios de información secundaria diferentes).

Los estudios de solapamiento se realizan mediante el cotejo de las listas de títulos o las referencias procesadas por cada uno de los productos estudiados y la extracción de los títulos o referencias con cualquier grado de multiplicación, es decir repetidos al menos en 2 o más de los productos o servicios analizados.

La aplicación de la técnica de solapamiento implica la comparación de los resultados de las tareas de evaluación y selección realizadas por diferentes servicios de información secundaria, que en general se escogen entre los de mayor adecuación y prestigio en las disciplinas estudiadas.

En esta comparación, se tratan de hallar los elementos que en el cubrimiento de la literatura resultan comunes a dos o más servicios de información secundaria. Los elementos más comúnmente utilizados son las publicaciones periódicas y las referencias.

Mediante esta técnica no se evalúan directamente las publicaciones y referencias sino que se comparan los resultados de las evaluaciones otorgadas a ellas por los sistemas que se estudian.

Un título que se seleccione de esta forma posee un efecto multiplicador en cuanto a su evaluación, porque en ésta participaron más especialistas, así como se emplearon más variables y procedimientos diferentes que en cualquier otro título que sea cubierto en forma única, es decir, por sólo uno de los productos analizados.

FUNDAMENTOS TEÓRICOS Y METODOLÓGICOS PARA LA REALIZACIÓN DE LOS ESTUDIOS DE SOLAPAMIENTO

El aspecto más importante que se impone esclarecer es el valor de las referencias y de los títulos únicos y solapados.

Cada servicio de información secundaria cubre una materia bajo un conjunto de criterios con los cuales pretende seleccionar aquellos títulos y trabajos con una mayor probabilidad de satisfacer las necesidades de sus usuarios.

Como cada comunidad de usuarios a la que se pretende servir es posible que difiera en algún grado de otras (por ejemplo, los clínicos de EE.UU. y Francia con 2 escuelas de medicina distintas o los médicos de países desarrollados y subdesarrollados), necesariamente aparecerán similitudes y diferencias en el procesamiento que se realice de la literatura internacional con respecto a un tema determinado en cada servicio de información.

Las semejanzas y divergencias de las necesidades e intereses de las comunidades de usuarios generan similitudes y diferencias en el procesamiento de la literatura mundial. Como consecuencia aparecerá un número indeterminado de referencias solapadas y únicas en cada producto de información secundaria que se estudie con respecto a los demás existentes en su misma categoría. Dichas similitudes y diferencias son el resultado de la selectividad del procesamiento, tanto de los títulos, como de los trabajos aparecidos en las distintas publicaciones. Aun cuando formalmente puede aparecer una publicación procesada en 2 o más bases de datos, el número de trabajos que de ella aparezcan en ambas bases puede ser diferente.

La aparición de una revista de carácter único o con un número mayor de referen-

cias puede ser el resultado de distintos factores como son:

- La publicación presenta una utilidad particular a la comunidad para la que se procesa.
- La base de datos realiza un cubrimiento exhaustivo de la materia a la que pertenece la revista.
- El idioma y el país de la revista se encuentran entre los que alcanzan el servicio de información secundaria, entre otros.

Es por éstas y otras razones, que la unicidad en el procesamiento no puede tomarse como sinónimo de un mayor alcance o cubrimiento de la literatura de una rama del saber por la base de datos analizada porque, en ciertas ocasiones, este cubrimiento específico obedece a la localidad de los intereses que representa el servicio o a la particularidad de los requerimientos de un contexto determinado.

Otros factores, como el interés de registrar toda la literatura que genera una región o país concreto o simplemente la intención del país creador de la base de datos de colocar en su producto la mayor cantidad de literatura nacional posible, aun a despecho de su utilidad o calidad, pueden invalidar el criterio de unicidad.

Con independencia de que en el plano teórico, las llamadas "referencias únicas" pueden ser de igual utilidad y calidad que las solapadas, estas últimas presentan un menor riesgo de ser el producto de las particularidades que caracterizan a cualquier base de datos o de las diferencias en su procesamiento. Una publicación solapada en general es una publicación más conocida, evaluada y, por tanto, con mejores probabilidades de presentar una utilidad y una calidad superior.

Las referencias únicas, sin embargo, pueden constituir referencias de gran utili-

dad, interés y calidad, cuando son el producto de un rastreo exhaustivo y sistemático de la literatura publicada en un área del conocimiento en el nivel mundial por parte de los creadores de la base de datos analizada, así como de una valoración exigente como la que requieren (aún en mayor grado) las revistas menos conocidas.

Cuando se cumplen estas condiciones, el número de referencias únicas puede utilizarse como un buen indicador para la valoración de las bases de datos.

La valoración de la utilidad y la calidad de las referencias únicas exige de la evaluación a ciegas, es decir sin el conocimiento por parte de los cuales las evalúan, de los autores, las instituciones, los países y los títulos de las publicaciones de las cuales proceden; por cuanto sin estas condiciones con frecuencia ocurre una valoración prejuiciada de las referencias.

DETERMINACIÓN DEL GRADO DE SOLAPAMIENTO DE LAS PUBLICACIONES

Como se ha expuesto antes, para la realización de un estudio del grado de solapamiento que presentan las revistas procesadas en varias bases de datos pueden utilizarse 2 entidades diferentes, los títulos de las publicaciones o las referencias de los trabajos publicados.

En la primera opción, el estudio resulta mucho más fácil y económico de realizar, por cuanto éste se limitará a determinar cuáles títulos aparecen registrados en más de una base de datos. La segunda alternativa producirá resultados mucho más precisos porque, como se explicó antes, la aparición de un mismo título en 2 o más bases de datos no condiciona la presencia por igual de todos sus trabajos en ambas bases, determinado esto por la selectividad del procesamiento que se realiza.

Sin embargo, la aplicación del procedimiento resultará mucho más costosa y compleja, sobre todo si dicho análisis se realiza de forma manual.

A continuación, se ejemplificará cómo es posible calcular el grado de solapamiento de un título mediante 2 variantes diferentes.

Este valor concede un peso superior a las referencias solapadas sobre las referencias de carácter único tal y como es deseable que suceda.

Por el contrario, si este valor se hallara de igual forma que para las referencias solapadas exclusivamente, el mayor peso de las referencias únicas en el título B, haría que éste alcanzara un valor más alto que el título A en la escala de valores, lo que no resulta adecuado con este enfoque.

Ejemplo 1. Cálculo del grado de solapamiento sin referencias únicas incluidas.

Grado de solapamiento = Σ categoría de la referencia x Número de referencias

Títulos	Referencias Solapadas	Referencias Categoría 3 (de triple solapamiento)	Referencias Categoría 2 (de doble solapamiento)
Revista A	7	3	4
Revista B	6	3	3

Grado de solapamiento:

de la revista A	de la revista B
$= 3 \times 3 + 2 \times 4$	$= 3 \times 3 + 2 \times 3$
$= 9 + 8$	$= 9 + 6$
$= 17$	$= 15$

Ejemplo 2. Cálculo del grado de solapamiento proporcional (variante del ejemplo 1) con referencias únicas incluidas.

Grado de solapamiento = Σ Categoría de la referencia x Proporción de referencias en la categoría

Títulos	Referencias recuperadas	Referencias solapadas	Distribución por categoría de las referencias		
			De la categoría 3 (3 puntos)	De la categoría 2 (2 puntos)	De la categoría 1 (única o de 1 punto)
Revista A	10	7	3	4	3
Revista B	12	6	3	3	6

Para este cálculo los valores de las categorías deben ser convertidos en proporciones, con el objetivo de mantener el peso de las referencias solapadas sobre las referencias únicas.

De aquí que para ambas revistas la distribución por categorías quede de la forma siguiente:

Revista A: $\frac{3}{10} = 0,33$; $\frac{4}{10} = 0,40$; $\frac{3}{10} = 0,33$

Revista B: $\frac{3}{12} = 0,25$; $\frac{3}{12} = 0,25$; $\frac{6}{12} = 0,50$

Grado de solapamiento (con referencias únicas incluidas)

Revista A	Revista B
$= 3 \times 0,33 + 2 \times 0,40 + 1 \times 0,33$	$= 3 \times 0,25 + 2 \times 0,25 + 1 \times 0,50$
$= 0,99 + 0,80 + 0,33$	$= 0,75 + 0,50 + 0,50$
$= 2,1$	$= 1,75$

Lo hasta aquí tratado, resume la primera opción metodológica, es decir, cuando se toma el criterio de un índice mayor de solapamiento como primario.

Ahora bien, en ciertas ocasiones, especialmente cuando se cumplen las condiciones requeridas para que la unicidad resulte un criterio válido y con el propósito de evaluar el cubrimiento de la literatura que realiza una base de datos en una materia determinada, resulta muy útil el empleo de una alternativa metodológica denominada *Índice de Meyer*.

EL ÍNDICE DE MEYER EN LA EVALUACIÓN DE LAS BASES DE DATOS Y LAS PUBLICACIONES PERIÓDICAS

El Índice de Meyer o índice relativo de peculiaridad se emplea para el estudio de bases de datos mediante la repetición en cada una de ellas de un conjunto de búsquedas predeterminadas.

Establece un peso básico para cada una de la bases de datos que resulta tanto más alto cuanto mayor sea el número de referencias únicas encontradas.

Bajo este criterio, las referencias únicas son las que toman mayor valor y las solapadas se dividen entre la unidad. Por ejemplo, una referencia de doble solapamiento adquiere valor de $0,5 (1/2) = 0,5$.

Mediante este índice, se evalúa la exhaustividad del cubrimiento de las bases de datos consideradas en una temática.

Para producir un número mayor de referencias únicas como resultado de una búsqueda en determinada temática es necesario procesar una cantidad mayor de publicaciones dentro del campo de conocimiento del que se trate, y por lo tanto, cubrir una cantidad mayor de títulos en forma única.

De manera general, resulta necesario para lograr una mayor exhaustividad en el cubrimiento de una materia, procesar un conjunto de títulos no considerados importantes internacionalmente en la temática (y esto no

Ejemplo 3. Cálculo del Índice de Meyer aplicado a las publicaciones.

$$\text{Índice Meyer} = \frac{\sum \text{Número de referencias} \times \text{Valor de la categoría}}{\text{Total de referencias recuperadas}}$$

$$\text{Revista A} = \frac{3 \times 0,3 + 4 \times 0,5 + 1 \times 3}{10}$$

$$= \frac{0,9 + 1,5 + 3}{10}$$

$$= \frac{5,9}{10} = 0,59$$

$$\text{Revista B} = \frac{3 \times 0,3 + 3 \times 0,5 + 1 \times 6}{12}$$

$$= \frac{0,9 + 1,5 + 6}{12}$$

$$= \frac{8,4}{12} = 0,70$$

(0,3 corresponde a un triple solapamiento y 0,5 a un doble solapamiento).

indica que en realidad no lo sean) para buscar referencias de carácter único. Exige, a su vez, una mayor disponibilidad de recursos por parte de la institución que lo realiza.

Su cálculo se realiza con la división del total de la sumatoria de los valores alcanzados por las referencias entre el total de referencias recuperadas. Su valor resultará tanto más alto cuanto más se acerque a la unidad.

Si se emplearan las referencias utilizadas en los títulos A y B del último ejemplo, se hallaría el índice de unicidad para cada revista de la forma siguiente:

Este valor concede mayor peso a las referencias únicas o de menor solapamiento sobre los de mayor grado de repetición entre las bases de datos analizadas. El cálculo del índice para las bases de datos se realiza de igual forma.

La selección de las bases de datos debe dirigirse a la obtención de aquellos productos de información secundaria más exhaustivos dentro de sus materias de procesamiento, siempre y cuando se cumplan las condiciones enunciadas.

Por su parte, el análisis de las publicaciones periódicas debe estar dirigido a determinar aquellas revistas que presentan un mayor grado de solapamiento o efecto multiplicador, entre los servicios de información secundaria que emplea o pretende utilizar la institución de información.

CONDICIONES PARA LA APLICACIÓN DE LA TÉCNICA DE SOLAPAMIENTO

Para elevar el valor de los resultados obtenidos en los estudios de solapamiento deben cumplirse las condiciones siguientes:

- Existencia de una comunidad temática entre los servicios de información secundaria analizados.
- Presencia de una correspondencia entre los atributos del procesamiento, por ejemplo, alcance temático, niveles de profundidad (exhaustivo o superficial)

y tipo (completo o selectivo), los períodos de actualización, la especificidad de los sistemas de recuperación, entre otros, y los requerimientos de las necesidades de la comunidad que se trata de servir.

- Preferencia de los servicios de información secundaria ubicados en zonas geográficas (países, regiones, continentes, etc.) diferentes.

En el estudio de los servicios de información secundaria, la utilización del solapamiento entre sistemas ubicados en países, regiones y continentes diferentes, elimina las preferencias que causan la cercanía geográfica, los idiomas, la cultura, etc.; razón por la cual se eleva la confiabilidad de los resultados en la determinación de las referencias y los títulos que son verdaderamente importantes.

La comparación entre continentes resulta conveniente cuando se trata de extraer un conjunto de publicaciones de interés mundial en la temática. Sin embargo, cuando se aplica en regiones y países diferentes, permite conocer el interés y la utilidad para regiones y países con características específicas y matices diversos.

Una variante en estos estudios -que no es una condición para los análisis de solapamiento- es la utilización de la categoría tiempo. En ella, se comparan las listas de títulos procesados en 2 momentos diferentes por una misma base de datos. Cuando se comparan listas cuyas fechas de edición se diferencian por un período corto de tiempo que puede oscilar entre 2 y 5 años, según la materia objeto de análisis, es posible obtener aquellas publicaciones periódicas que presentan una mayor estabilidad dentro de la materia.

Las publicaciones periódicas halladas mediante esta variante, se encuentran profundamente "enraizadas" en el caudal informacional de la especialidad, "aisladas de la superficie ebullente que caracteriza al mundo editorial en un momento específico en la mayoría de las disciplinas científicas".

PRINCIPALES VENTAJAS Y DESVENTAJAS DE LOS ESTUDIOS DE SOLAPAMIENTO

Entre las ventajas de su aplicación se destacan:

- Sencillez y facilidad de aplicación.
- Carácter eminentemente económico.
- Obtención de publicaciones con un mayor efecto multiplicador.
- Posibilidad de determinar el grado de exhaustividad que presentan las bases de datos.

Dos de las incógnitas que resuelven estos análisis son: cuáles títulos debe adquirir una institución para respaldar con mayor efectividad y eficiencia al conjunto de bases de datos con las que se presta servicio y cuáles bases de datos nuevas deben adquirirse para lograr una extensión significativa en el cubrimiento de la literatura que realiza el sistema en su área de conocimiento.

En ambos casos, se logra un perfeccionamiento de la utilización de los recur-

sos de la institución con un aumento significativo en la efectividad de sus servicios.

Algunas de las dificultades para realizar estos análisis son:

- Los servicios de información secundaria cubren algunas materias de forma exhaustiva y otras pobremente, bien porque son muy selectivos o porque la materia no es de su interés.
- Las temáticas nuevas o interdisciplinarias no siempre son bien cubiertas.
- La existencia de una serie de preferencias geográficas, idiomáticas, culturales, políticas y de otros tipos que caracterizan individualmente a cada base de datos.

Los resultados obtenidos mediante los estudios de solapamiento deben considerarse bajo la óptica de éstos y otros factores, que pueden influir de manera sensible en su conformación final.

Abstract

Overlapping studies in the selection of serials/ and data bases

Selecting serial publications and databases turns out to be a very difficult task, especially in the biomedical sector due to the accelerated increase of information sources and means in this area of knowledge. In that sense, the conducting of infommetrical researches is a current trend. Within this, overlapping studies represent an useful way to determine the degree of processing that a number of publications of an specific item receive in different databases and in turn allow to find the level of exhaustivity that these present in covering the literature of an area of knowledge. The theoretical and methodological grounds of the overlapping studies are exposed and also some practical ways for its implementation in assessing the titles of serials such as the databases, as well as its advantages and limitations.

Subject headings: SERIALS; DATA BASES; SELECTION; LAPPING; INFORMETRICS/ trends.

REFERENCIAS BIBLIOGRÁFICAS

1. Morales Morejón M comp. Introducción a la informetría. Glosario. La Habana: Sociedad Cubana de Información Científica y Tecnológica. Instituto de Documentación e Información Científica y Tecnología, 1988:107.

Recibido: 25 de junio de 1999.

Aprobado: 14 de julio de 1999.

Lic. Rubén Cañedo Andalia. Centro Nacional de Información de Ciencias Médicas. Calle E No. 452 e/19 y 21, El Vedado, Ciudad de La Habana, Cuba.