

# Portal de indicadores bibliométricos BIBLAT

Gonzalo Reséndiz Cansino, Antonio Sánchez Pereyra,  
Mauricio Fabián Sánchez R., Isela García Bravo

## RESUMEN

Los indicadores bibliométricos y cientométricos han adquirido mayor notoriedad a raíz del surgimiento, relativamente reciente, de *rankings* enfocados en la evaluación y comparación del desempeño de las universidades a nivel mundial. A pesar de sus diferencias metodológicas, los *rankings* internacionales se basan en la información que proveen las bases de datos *Web of Science* y *Scopus*. Sin embargo, dichos *rankings* y bases de datos no son lo suficientemente representativos de la ciencia desarrollada en América Latina y el Caribe. Esta situación resulta desfavorable para el reconocimiento de la producción científica regional que es publicada en revistas que no son indizadas en las principales bases de datos internacionales. En este trabajo se presentan los objetivos, la funcionalidad y la plataforma tecnológica de BIBLAT, programa enfocado a la generación de indicadores de la ciencia regional. El portal BIBLAT es respaldado por CLASE y PERIÓDICA, dos de las bases de datos de publicaciones periódicas más antiguas y representativas de América Latina y el Caribe. Los indicadores generados por BIBLAT están basados en la productividad medida en términos del número de documentos publicados por autor, institución/universidad de afiliación del autor y su nacionalidad, revistas y su nacionalidad, áreas del conocimiento y colaboración intra y extra regional.

**PALABRAS CLAVE:** Indicadores bibliométricos, América Latina, revistas científicas, protocolo HTTP, HTML, Javascript, URL, OPAC, base de datos, interfaz, PHP, MySQL, cliente-servidor, software libre, CLASE, PERIÓDICA, ALEPH.

## Abstract

Recently, the relevance of bibliometric and scientometric indicators has been increased by some of new rankings that compare and evaluate the performance of all world universities. Notwithstanding the methodological differences between these rankings all of them use the data from *Web of Science* and *Scopus*. However, these rankings and databases are not sufficiently representative of the Latin American and the Caribbean science. This situation is adverse for the recognition of the regional scientific production published in journals that have not been indexed in the major journal databases. This paper presents the objectives, functionality and informatic platform of BIBLAT, a project focused in the generation of regional science indicators. BIBLAT portal is supported by CLASE and PERIODICA, both of the oldest and most representative journal databases from Latin America and the Caribbean. The indicators generated by BIBLAT are based on measurement of productivity in terms of the number of articles published by author, author's institutions/universities and their nationality, journals and their nationality, knowledge areas, and intra and extra regional collaboration.

**KEYWORDS:** Bibliometric indicators, Latin America, scientific journals, HTTP Protocol, HTML, Javascript, URL, OPAC, databases, interface, PHP, MySQL, client-server, free software, CLASE, PERIÓDICA, ALEPH.

## Introducción

**E**l portal BIBLAT (<http://biblat.unam.mx>) está integrado tomando como referencia las bases de datos CLASE y PERIÓDICA, y ha sido concebido como una fuente de información que proporciona datos bibliométricos sobre la producción científica publicada en revistas latinoamericanas y del Caribe indizadas en las bases de datos CLASE (CITAS LATINOAMERICANAS EN CIENCIAS SOCIALES Y HUMANIDADES) y PERIÓDICA (ÍNDICE DE REVISTAS LATINOAMERICANAS EN CIENCIAS). El objetivo de BIBLAT es ofrecer información cuantitativa para conocer y comparar la aportación científica de los autores, instituciones y publicaciones involucradas en la investigación y la divulgación del conocimiento en América Latina y el Caribe. Para tal fin, BIBLAT conjunta y sistematiza la información contenida en CLASE y PERIÓDICA, las bases de datos bibliográficas con el conjunto más representativo e incluyente de revistas académicas publicadas en América Latina y el Caribe (más de 3,000 revistas y más de 600,000 registros bibliográficos entre ambas bases). Con este sustento, BIBLAT permite la obtención de indicadores sobre artículos y documentos publicados en las revistas de la región, de acuerdo con las siguientes categorías: producción de artículos por autor, por institución de afiliación del autor, por país de estas instituciones, así como por revista y país de la revista; además, BIBLAT genera información sobre las áreas temáticas (o disciplinas del conocimiento) así como la colaboración intra y extra regional identificada a través de los trabajos en coautoría a nivel de instituciones de afiliación de los autores y de los respectivos países de estas instituciones.

## Antecedentes

La ciencia generada en América Latina, como la del resto de países pertenecientes al Tercer Mundo, ha sido identificada como “ciencia periférica”, e incluso con una caracterización más severa aún, la de “ciencia perdida”, debido a diversos factores, entre ellos los bajos niveles de impacto y difusión de sus revistas científicas (Gibbs 1995). En contrapartida, la necesidad de impulsar la visibilidad y el reconocimiento nacional e internacional de estas publicaciones ha propiciado numerosas iniciativas centradas en la creación de sistemas de información y

bases de datos especializadas en revistas de la región. CLASE y PERIÓDICA se encuentran entre las bases de datos de mayor pervivencia (CLASE inició en 1976 y PERIÓDICA en 1978) y cobertura temática, ya que ambas bases son multidisciplinarias. Concebidas originalmente en función de proveer servicios de información a usuarios, así como para diseminar el conocimiento científico y académico publicado en las revistas nacionales y de la región en su conjunto, CLASE y PERIÓDICA han tenido que responder a otro tipo de demanda de información consistente en indicadores bibliométricos y cienciométricos, la cual es solicitada por un sector muy especializado de usuarios dedicados al estudio del comportamiento de las comunidades científicas y por instancias y autoridades encargadas de la política científica y tecnológica.

A nivel mundial, existen dos bases de datos que se desempeñan como el referente para la medición de la producción científica cuantificada en número de artículos publicados en revistas académicas reconocidas, así como del impacto de éstas en la comunidad científica. Estas bases de datos son *Web of Science* (ISI Thomson Reuters) y *Scopus* (Elsevier). La relevancia de los indicadores bibliométricos se ha visto acrecentada últimamente por el surgimiento de diversos *rankings* internacionales que comparan y jerarquizan el desempeño de las universidades del mundo, como lo son el *Academic Ranking of World Universities* (ARWU), desarrollado por la universidad china Shanghai Jiao Tong, el *Times Higher Education* (THE) realizado por la revista inglesa *Times*, los *rankings* e indicadores ofrecidos por el proyecto español *scimago* (el cual está basado exclusivamente en la información proporcionada por *Scopus*), así como los diversos *rankings* basados en la webometría. Diferenciados unos y otros por las variables de sus metodologías particulares, todos comparten la ponderación del impacto de la investigación de las universidades reflejado en los análisis de citación ofrecidos por *Web of Science* y *Scopus* (y de Google Scholar en el caso de la webometría).

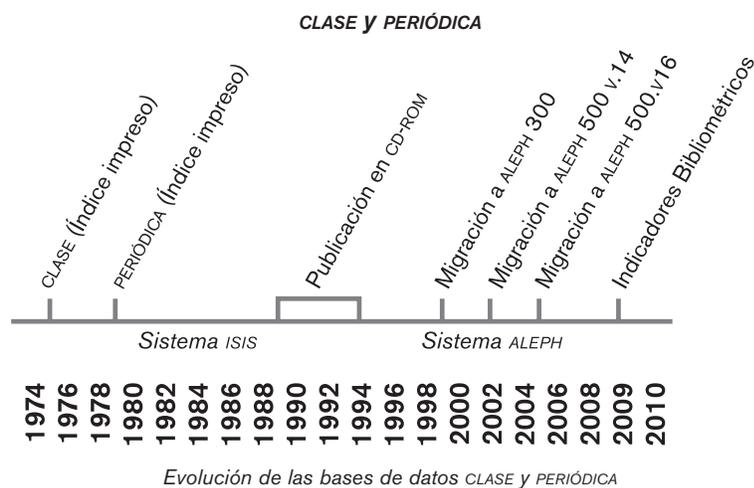
Además de los sesgos cuantitativos que conlleva este tipo de evaluación del desempeño de una institución y de los países de dichas instituciones, para el caso de América Latina y el Caribe (lo cual es extensivo al conjunto iberoamericano) la problemática se ve agravada por el

escaso nivel de integración que históricamente han logrado las publicaciones de la región en el *Web of Science*, lo cual repercute desfavorablemente en la representatividad y el reconocimiento de la producción científica regional. Ciertamente, dicha situación ha sido atemperada con el tiempo y encuentra perspectivas favorables tanto por el hecho de que ha aumentado el número de investigadores de la región que publican en revistas extranjeras, como por el programa de inclusión masiva de las denominadas revistas "regionales" lanzado por *Web of Science* desde 2006, el cual ha redundado en la inclusión de más títulos latinoamericanos. Por otra parte, *Scopus*, sistema que salió al mercado de manera relativamente reciente (2004), tuvo desde sus orígenes una mayor apertura hacia la inclusión de títulos de la región.

No obstante lo anterior, la especialización en un mayor número de títulos de la región geográfica y la masa crítica de información acumulada a través de más de 30 años de trayectoria de CLASE y PERIÓDICA convierten a BIBLAT en una fuente alternativa y complementaria a *Web of Science* y *Scopus*, lo que permite ofrecer una aproximación más representativa de la publicación científica latinoamericana, contribuyendo así a reflejar el volumen, la dinámica y los análisis comparativos sobre la actividad científica en aquellas áreas de estudio cuya comunicación no transcurre a través de las publicaciones de alcance internacional, publicadas en idioma inglés, sino a través de las revistas locales y escritas en español y portugués, como es el caso de las ciencias aplicadas (agronomía y medicina, por ejemplo) y, de manera destacada, las ciencias sociales y humanidades. Debe mencionarse que los indicadores generados por BIBLAT no están basados en la citación (número de citas recibidas por documento) sino en la cuantificación de la productividad en términos del número de documentos por autor, institución, revista, país de la institución y coautorías interinstitucionales y entre países.

Por su parte, CLASE y PERIÓDICA han evolucionado conforme a las tendencias de las tecnologías de la información y la comunicación a nivel internacional. Primero, surgieron como índices impresos a través del sistema *CDS/ISIS*<sup>1</sup>, cu-

yas versiones posteriores fueron *MINISIS* y *MicroISIS*; actualmente residen en el sistema *ALEPH* en la versión 500.16 y cuentan con una interfaz web de consulta (CLASE <http://clase.unam.mx>; PERIÓDICA <http://periodica.unam.mx>).



A continuación se muestra la cobertura por áreas del conocimiento para cada una de las bases:

| CLASE                                       |            |                            |
|---|------------|----------------------------|
| Administración y contaduría                 | Demografía | Historia                   |
| Antropología                                | Derecho    | Literatura y lingüística   |
| Arte  | Economía   | Psicología                 |
| Bibliotecología y ciencia de la información | Educación  | Relaciones internacionales |
| Ciencia política                            | Filosofía  | Religión                   |
| Ciencias de la comunicación                 | Geografía  | Sociología                 |

| PERIÓDICA                  |                     |                         |
|----------------------------|---------------------|-------------------------|
| Agrocencias                | Física y astronomía | Matemáticas             |
| Biología                   | Geociencias         | Medicina                |
| Ciencias de la Computación | Ingeniería          | Química                 |
| Ciencia y tecnología       |                     | Veterinaria y zootecnia |

<sup>1</sup> ISIS (Integrated Set of Information System), desarrollado por la Oficina Internacional del Trabajo.

Actualmente estas bases cuentan con una amplia cobertura de publicaciones indizadas, tal como se muestra en la siguiente tabla:

| Base de datos | Revistas | Artículos |
|---------------|----------|-----------|
| CLASE         | 1,649    | 315,579   |
| PERIÓDICA     | 1,514    | 326,399   |

Fuente: Bases de datos CLASE y PERIÓDICA (información al 30 de Septiembre de 2010)

CLASE y PERIÓDICA también son fuente de información para los siguientes sistemas de información y proyectos enfocados en la generación de indicadores bibliométricos:

- Online Computer Library Center (*OCLC*): desde 2003 CLASE y PERIÓDICA forman parte del conjunto de bases de datos del servicio *FirstSearch* que ofrece esta compañía estadounidense.
- Red Iberoamericana de Indicadores de Ciencia y Tecnología (*RICYT*): proyecto coordinado desde Argentina.
- Estudio Comparativo de las Universidades Mexicanas (*ECUM*)

### Contexto sobre indicadores bibliométricos

CLASE y PERIÓDICA han sido incorporadas como fuentes de información en dos proyectos enfocados en la evaluación del desempeño científico y académico basado en indicadores cuantitativos: la Red de Indicadores en Ciencia y Tecnología (*RICYT* – <http://www.ricyt.org>) y el Estudio Comparativo de Universidades Mexicanas (*ECUM* – <http://www.ecum.unam.mx>), portal recientemente desarrollado (2010) por la UNAM, el cual está especializado en reportes comparativos sobre la producción científica de las universidades y centros de investigación del país. En la siguiente imagen se puede apreciar una hoja de resultados de búsqueda en EXECUM, que es el explorador o herramienta informática que permite realizar búsquedas y visualizar los

datos de ECUM; la información extraída de CLASE y PERIÓDICA está organizada en los siguientes rubros: Producción institucional, la cual se subdivide, a su vez, en Número de artículos totales (esto es, la producción de artículos y documentos de una universidad mexicana, en este ejemplo la UNAM), y el Porcentaje nacional (que es el porcentaje de artículos y documentos publicados por autores de una universidad en relación con el total de la producción mexicana); Análisis por autor, subdividido en Número de artículos en los que participa como primer autor un académico de la universidad referida, el Porcentaje nacional (que es el porcentaje de artículos y documentos en los que hay participación como primer autor por parte de un académico de la universidad referida respecto de la producción mexicana) y el Porcentaje institucional (que es el porcentaje en el que hay participación como primer autor respecto de la producción de la misma institución); el tercer rubro son las Colaboraciones (o coautorías), subdivididas en las Colaboraciones totales (número de artículos y documentos producidos en colaboración) y Porcentaje nacional (porcentaje de artículos y documentos de la universidad referida respecto de la producción mexicana realizada en colaboración); finalmente, figuran la producción de Colaboraciones nacionales totales (número de artículos y documentos elaborados en coautoría exclusivamente nacional), así como los porcentajes en relación con el total nacional y al total institucional, y Colaboraciones internacionales (número de artículos y documentos elaborados en coautoría con al menos un autor extranjero), así como el Porcentaje nacional (porcentaje de artículos y documentos respecto de la producción mexicana realizada en coautoría internacional) y el Porcentaje institucional (porcentaje de artículos y documentos respecto de la producción institucional realizada en coautoría internacional).

Por su especialización regional, BIBLAT puede ser comparado con el Ranking Iberoamericano SIR<sup>2</sup> elaborado recientemente por el grupo de investigación *scimago* (Mayo 2010). Dicho informe incluye 607 instituciones iberoamericanas de Educación Superior que publicaron algún artículo científico durante el periodo 2003-2008 y haya sido indizado en *Scopus*, base de datos con una

<sup>2</sup> *Ranking Iberoamericano SIR 2010* [en línea]. Scimago Institutions Rankings. <[http://www.scimagoir.com/pdf/ranking\\_iberamericano\\_2010.pdf](http://www.scimagoir.com/pdf/ranking_iberamericano_2010.pdf)> [Consulta: 27 agosto 2010].

Univ. Seleccionadas    Primeras 20    Selección Personalizada

DGEI >> Estudio Comparativo >> Elaborar Reporte >> Participación Reportada (2008)

Universidades Seleccionadas CLASE y PERIÓDICA-(Artículos),2008

Metodología    Guía de usuario

Glosario de términos    Exportar a Excel

Para conocer la descripción del contenido de cada columna coloque el puntero del mouse encima del rótulo correspondiente

| No.                             | Institución                             | CLASE y PERIÓDICA (Artículos) |        |         |                    |        |         |                |        |         |          |        |         |               |        |         |
|---------------------------------|---|-------------------------------|--------|---------|--------------------|--------|---------|----------------|--------|---------|----------|--------|---------|---------------|--------|---------|
|                                 |   | Producción Institucional      |        |         | Análisis por autor |        |         | Colaboraciones |        |         |          |        |         |               |        |         |
|                                 |   | Total                         |        |         | Total              |        |         | Totales        |        |         | Nacional |        |         | Internacional |        |         |
|                                 |   | Núm.                          | % Nat. | % Inst. | Núm.               | % Nat. | % Inst. | Núm.           | % Nat. | % Inst. | Núm.     | % Nat. | % Inst. | Núm.          | % Nat. | % Inst. |
| 1                               | UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO | 499                           | 23.5   | 422     | 19.8               | 84.6   | 161     | 23.8           | 32.3   | 143     | 24       | 28.7   | 32      | 26.2          | 6.4    |         |
| <b>Totales 58 Universidades</b> |   | 1295                          | 100    | 1295    | 100                | 100    | 453     | 100            | 35     | 404     | 100      | 31.2   | 77      | 100           | 5.9    |         |
| <b>Totales Nacionales</b>       |   | 2126                          | 100    | 2126    | 100                | 100    | 677     | 100            | 31.8   | 597     | 100      | 28.1   | 122     | 100           | 5.7    |         |

ND: Dato no disponible

Notas:

- Se refieren al total de documentos / artículos en los que participa al menos un autor adscrito a una institución mexicana. Por la autoría múltiple, el número de documentos o artículos es inferior al total de participaciones institucionales. Por esta razón, la suma de las participaciones de cada institución es mayor que el total de documentos o artículos que la cifra que aparece en la tabla.
- Se refieren a porcentajes, sumatorias y promedios nacionales calculados con base en el total de documentos / artículos en los que participa al menos un autor adscrito a una institución mexicana.

Fuentes de información: ISI WoK, Clase y Periódica, CONACYT, SEP

Fuente: Portal ECUM-EXECUM

cobertura de más de 17,000 títulos de revistas científicas de todo el mundo. El *ranking* consta de cuatro indicadores: 1) producción científica, o número de artículos publicados por las universidades; 2) colaboración internacional, medida como la proporción (*ratio*) de artículos que publica una universidad en colaboración con universidades de otros países, 3) calidad científica promedio, medida como la proporción (*ratio*) entre la citación que recibe una universidad y la citación media mundial y, 4) porcentaje de publicaciones en el 25% de las “mejores revistas del mundo”, ordenadas utilizando el indicador de relevancia científica “*Scimago Journal Ranking*” (SJR, este indicador incluye un algoritmo para calcular el factor de impacto de la revista que, entre otras virtudes, pondera la calidad de la fuente de cada cita). En términos comparativos, la cobertura de *scimago* es iberoamericana, esto es, incluye también a España y Portugal y no solamente América Latina y el Caribe como es el caso de BIBLAT. Por otra parte, BIBLAT ofrece la categoría de indicadores por producción científica y colaboración inter-

nacional de *scimago*, pero no los indicadores de calidad científica promedio y porcentaje de publicaciones, ya que estos indicadores están basados en análisis de citación. Mientras que el reporte de *scimago* está basado en la cobertura de más de 17,000 revistas de todo el mundo del índice *Scopus*, la especialización regional de BIBLAT se manifiesta en el dato de que están identificadas miles de instituciones y centros de estudio de la región, y no solamente las 607 instituciones a las que está circunscrito el informe de *scimago*, lo que otorga visibilidad a la producción académica del variado y multifacético entramado sistema científico y educativo de la región.

Además, los reportes de BIBLAT incluyen también producción por autor, por revista y país de la revista; por otro lado, los reportes de colaboración (coautorías) pueden realizarse de acuerdo con la selección del usuario, ya sea entre instituciones como entre países de dichas instituciones. De gran relevancia es también la capacidad de BIBLAT para generar reportes por área del cono-



En la siguiente imagen se observa la comparación entre un reporte de producción por institución de BIBLAT (seleccionando PERIÓDICA, base de datos fuente de BIBLAT) y el Ranking Iberoamericano SIR de *scimago*. Como se ha mencionado, entre las diferencias básicas se encuentran la cobertura y la especialización geográficas: el Ranking SIR incluye instituciones españolas y portuguesas ya que su cobertura es iberoamericana, BIBLAT reporta instituciones que no figuran en el informe del Ranking SIR (subrayadas), ya

que contiene una cobertura de revistas de América Latina y el Caribe más especializada que *Scopus* (base de datos fuente del Ranking SIR).

### Objetivo del proyecto

La finalidad de BIBLAT es generar indicadores bibliométricos a partir de las bases de datos de CLASE y PERIÓDICA, enfocados a la productividad académica de autores e instituciones de educación superior y de investigación.

| IBE* | LAC* | Institución  | País | PC     | CI    | CCP  | IQ    |
|------|------|--|------|--------|-------|------|-------|
| 1    | 1    | Universidade de Sao Paulo                              | BRA  | 37.352 | 24,81 | 0,81 | 40,35 |
| 2    | 2    | Universidad Nacional Autonoma de Mexico                | MEX  | 17.395 | 39,17 | 0,80 | 48,59 |
| 3    | 3    | Universidade Estadual de Campinas                      | BRA  | 14.913 | 21,47 | 0,81 | 38,18 |
| 4    |      | Universitat de Barcelona                               | ESP  | 14.742 | 41,64 | 1,41 | 62,16 |
| 5    |      | Universidad Complutense de Madrid                      | ESP  | 12.315 | 32,81 | 1,10 | 52,62 |
| 6    | 4    | Universidade Estadual Paulista Julio de Mesquita Filho | BRA  | 12.270 | 16,8  | 0,63 | 31,73 |
| 7    | 5    | Universidade Federal do Rio de Janeiro                 | BRA  | 12.133 | 26,67 | 0,80 | 39,01 |
| 8    |      | Universitat Autonoma de Barcelona                      | ESP  | 10.911 | 38,98 | 1,37 | 58,15 |

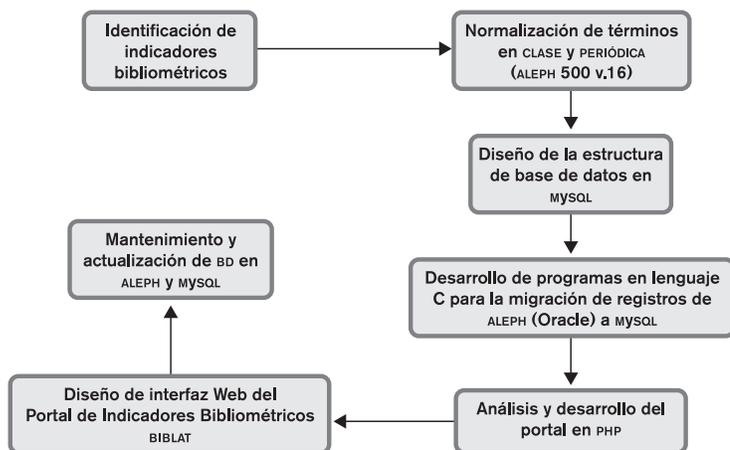
**SCImago SIR 2010**

**BIBLAT (PERIÓDICA)**

| Indicadores por institución de afiliación del autor |  |                          |
|---|--|--------------------------|
| No.   | Institución  | Documentos               |
| 1   | Universidad Nacional Autónoma de México                                      | País de la revista 35325 |
| 2   | Universidade de Sao Paulo  | País de la revista 18666 |
| 3   | Universidade Estadual Paulista "Julio de Mesquita Filho"                     | País de la revista 10063 |
| 4   | Instituto Mexicano del Seguro Social   | País de la revista 8337  |
| 5   | Universidad Autónoma Metropolitana   | País de la revista 7485  |
| 6   | Universidad de Chile   | País de la revista 5685  |
| 7   | Instituto Politécnico Nacional   | País de la revista 5436  |
| 8   | Universidade Federal do Rio de Janeiro                                       | País de la revista 5040  |
| 9   | Universidade Estadual de Campinas  | País de la revista 4482  |
| 10  | Universidade Federal de Minas Gerais   | País de la revista 4396  |
| 11  | Universidad de Buenos Aires  | País de la revista 4350  |
| 12  | Universidad Central de Venezuela   | País de la revista 3943  |
| 13  | Universidad de La Habana   | País de la revista 3630  |
| 14  | Empresa Brasileira de Pesquisa Agropecuária                                  | País de la revista 3552  |
| 15  | Universidade Federal do Rio Grande do Sul                                    | País de la revista 3527  |
| 16  | Universidade Federal de Vicosa   | País de la revista 3430  |
| 17  | Universidad de Costa Rica  | País de la revista 3359  |
| 18  | Universidad Nacional de Colombia   | País de la revista 3302  |
| 19  | Secretaría de Salud  | País de la revista 3283  |
| 20  | Universidade Federal do Parana   | País de la revista 3282  |
| 21  | Universidad de los Andes   | País de la revista 3263  |
| 22  | El Colegio de México   | País de la revista 3217  |
| 23  | Universidad Católica de Chile  | País de la revista 3175  |
| 24  | Universidad del Zulia  | País de la revista 2798  |
| 25  | Universidade Estadual de Maringa   | País de la revista 2700  |
| 26  | Universidad de Antioquia   | País de la revista 2461  |
| 27  | Universidad de Guadalajara   | País de la revista 2450  |
| 28  | Instituto Superior Politécnico "José Antonio Echeverría"                     | País de la revista 2361  |
| 29  | Colegio de Post-graduados  | País de la revista 2284  |
| 30  | Universidade de Brasilia   | País de la revista 2130  |
| 31  | Instituto Nacional de Antropología e Historia                                | País de la revista 2058  |
| 32  | Universidad Iberoamericana   | País de la revista 2036  |
| 33  | Universidad Austral de Chile   | País de la revista 1901  |
| 34  | Academia de Ciencias de Cuba   | País de la revista 1886  |
| 35  | Universidad de Concepción  | País de la revista 1881  |
| 36  | Universidade Estadual de Londrina  | País de la revista 1847  |
| 37  | Universidad Autónoma de Nuevo León   | País de la revista 1789  |
| 38  | Instituto de Seguridad y Servicios Sociales para los Trabajadores del Estado | País de la revista 1768  |
| 39  | Universidad Nacional de La Plata   | País de la revista 1768  |

## Metodología para determinar el modelo y su implementación

Como parte del desarrollo de este portal de indicadores bibliométricos se consideró la siguiente metodología:



Proceso de desarrollo del Portal de Indicadores Bibliométricos

BIBLAT

### Identificación de indicadores bibliométricos

En esta fase se definió el tipo de información que podrían ofrecer ambas bases de datos partiendo del objetivo principal de este proyecto, que es proporcionar indicadores de productividad académica tales como:

- Publicaciones por autor.
- Publicaciones por institución.
- Publicaciones por país de la institución de afiliación del autor.
- Publicaciones por revista.
- Publicaciones por país de la revista.
- Publicaciones por año de publicación.
- Publicaciones por área del conocimiento o disciplina.
- Coautorías entre países e instituciones.

### Normalización de términos en CLASE y PERIÓDICA

BIBLAT proporciona información sobre la adscripción institucional (o lugar de trabajo) de los autores de los

artículos o documentos publicados en las revistas indizadas en CLASE y PERIÓDICA, lo cual permite generar reportes cuantitativos sobre la producción de artículos y otro tipo de documentos considerando a las instituciones, organismos o centros de investigación en los que los autores desempeñan su actividad académica. Asimismo, BIBLAT genera reportes cuantitativos sobre la colaboración interinstitucional, tanto a nivel intrarregional (países de la región) como extrarregional (países externos a la región), misma que es identificada a través de los trabajos en coautoría. La generación de estos reportes es posible dado que CLASE y PERIÓDICA compilan –desde el inicio de ambas bases de datos– la información sobre la adscripción institucional del autor. Sin embargo, es necesario considerar la evolución de los criterios adoptados por dichas bases en la indización de esta información, dado que de ello depende su representación estadística. Hasta 1998 se registraban hasta tres niveles jerárquicos de cada institución: la institución y sus dependencias y subdependencias; a partir de 1998, sólo se codifican dos niveles: institución y dependencia, incluyendo ciudad, estado (o división político-administrativa) y país. Por otra parte, hasta 1987 sólo se registraba la información de adscripción institucional del primer autor; a partir de 1988, CLASE y PERIÓDICA registran todas las instituciones diferentes que aparecen en el documento. A partir de junio de 2009, se implementa una nueva política para la codificación de esta información, consistente en el registro de todas las instituciones que aparecen en el documento.

El sistema de gestión de bibliotecas ALEPH 500 ha sido de gran utilidad para el desarrollo de este proyecto, ya que al soportar el formato MARC facilita los procesos de normalización de la base de datos. La siguiente imagen proporciona dos diferentes vistas de un registro bibliográfico, por un lado la codificación en formato MARC indicando las distintas etiquetas y subcampos que conforman un registro, y por otra parte la vista para el usuario final a través del web OPAC.

Visualización de un registro bibliográfico en ALEPH 500v.16



Interfase Gráfica del Usuario (GUI)



Web OPAC

Otra ventaja de este formato es que proporciona compatibilidad con otros sistemas de información, lo que amplía las posibilidades de comunicación. El sistema ALEPH 500v.16 ha permitido llevar a cabo las tareas de normalización a través de herramientas propias del sistema, aunque en algunas ocasiones se han desarrollado programas para efectuar modificaciones masivas en toda la base bajo ciertos criterios de actualización. Así mismo, es importante señalar que este software permite compartir registros bibliográficos basándose en el protocolo de recuperación de información Z39.50.

**Diseño de la estructura de base de datos**

Dependiendo de los requerimientos de una base de datos el diseño de ésta puede resultar complejo, pero con simples reglas se puede tener una estructura funcional y eficiente; para este proyecto se consideraron los siguientes aspectos:

- Tipo de información.
- Cantidad de información.
- Velocidad de acceso.
- Facilidad para extraer información.
- Características del manejador de base de datos.

Otros aspectos importantes que se relacionan con el diseño de las bases de datos y que se tomaron en cuenta para este proyecto fueron los siguientes:

- Evitar la redundancia en los datos, para tener una recuperación rápida y eficiente.
- Diseño de tablas que modelen los registros que son almacenados.
- Almacenamiento de información exclusivamente necesaria, aspecto considerado también desde el punto de vista del desarrollo de la aplicación.
- Generación de índices para mejorar el rendimiento de las consultas en la base de datos.

Las principales entidades que integran esta base de datos son:

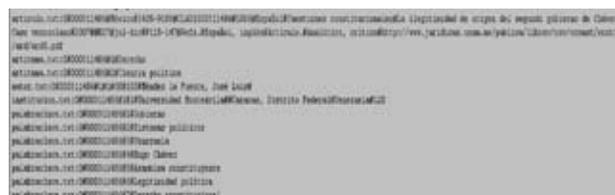
| Entidad              | Descripción   |
|----------------------|---|
| Artículo             | Contiene todos los documentos indizados en CLASE y PERIÓDICA, identificados a través del "ID" de la base (con datos tales como: idioma, título, año de publicación, volumen, número, páginas, resumen, enlace al texto completo, etcétera). |
| Autor                | Incluye a todos los Autores que participaron en la publicación del documento y un "ID" que identifica a la institución a la que están afiliados.  |
| Institución          | Hace referencia a la institución de afiliación para cada uno de los autores, incluyendo el país, ciudad y departamento en que colaboran.  |
| Palabra clave        | Contiene las palabras clave de ambas bases de datos.  |
| Disciplina artículo  | Hace referencia a la disciplina a la que pertenece el documento.  |
| Keyword              | Incluye las palabras clave en inglés de la base de datos de PERIÓDICA.  |
| Disciplina           | Es el catálogo de disciplinas de ambas bases de datos.  |
| Revistas-disciplinas | Es la clasificación de las revistas indizadas en ambas bases de datos por la disciplina a la que pertenecen.  |

**Desarrollo de programas en lenguaje C para la migración de registros de ALEPH (Oracle) a MySQL**

La ventaja de almacenar la información en formato ALEPH hace posible que los datos puedan migrarse o transferirse a otra plataforma de una manera fácil y sencilla, como se detalla a continuación:

1. Con los datos bien delimitados en cuanto al uso de etiquetas y subcampos MARC fue posible determinar el tipo de información que sería almacenada en cada una de las tablas de la base de datos.
2. Los procesos para recuperación de información que ofrece el sistema ALEPH 500v.16 (p\_ret\_01 y p\_print\_03) facilitaron la tarea para importar datos a otra plataforma.

3. El lenguaje de programación C nos ha facilitado la manipulación de los registros en formato ALEPH, de tal manera que se pueden crear archivos de texto plano delimitados por un caracter (ejemplo "#") para ser exportados a otras plataformas como MySQL, MS-Access, Excel, PostgreSQL, etcétera.



Archivos de texto plano delimitados por el caracter #



Registro secuencial bibliográfico en formato ALEPH obtenido a través del proceso p\_print\_03 de ALEPH 500v.16

**Análisis y desarrollo del Portal de Indicadores Bibliométricos BIBLAT**

El desarrollo en software libre es una excelente alternativa para la implementación de herramientas, además de que pueden ser portables fácilmente a otras plataformas con los mismos requerimientos de software; este proyecto ha sido desarrollado bajo la siguiente plataforma:

**Características del hardware - software**

- Servidor: M4000, 4 procesadores SUN SPARC64 VI a 2.15 GHZ.

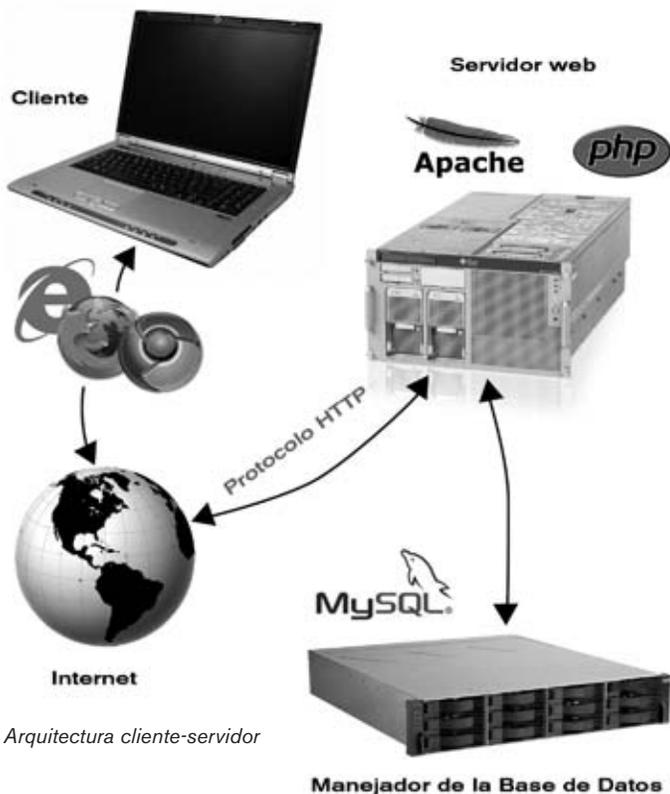
- Memoria RAM: 32 GB.
- Sistema Operativo: Solaris 10.
- Servidor Web: Apache Server.
- Manejador de Bases de Datos: MySQL 5.0.

#### Lenguajes de programación:

- HTML (HyperText Markup Language).
- PHP 5.3.
- Javascript.
- Lenguaje C.
- Lenguaje SQL (Structured Query Language).

#### Arquitectura Cliente-Servidor:

El Protocolo de Transferencia de Hipertexto (HTTP) permite intercambiar información bajo la arquitectura cliente-servidor, este protocolo establece la forma en que será transmitida la información a través de la red. Un cliente web (visualizador) es una aplicación capaz de enviar una solicitud a cualquier servidor web, esta aplicación podría ser Internet Explorer, Mozilla Firefox, Netscape, entre otras. Por otra parte, el servidor web recibe la solicitud del cliente y retorna una respuesta; esto permite que el cliente funcione de manera independiente del servidor web y se racionalicen recursos en la conexión, tal y como se representa en la siguiente gráfica:



Arquitectura cliente-servidor

Cabe mencionar algunas de las características del manejador de bases de datos denominado MySQL<sup>3</sup>:

- Es un manejador de bases de datos relacional y multiusuario.
- Es software libre en un esquema de licenciamiento dual, es decir que un producto derivado de éste puede ser licenciado para su uso comercial adquiriendo la licencia comercial de MySQL.
- Soporta el lenguaje estándar para bases de datos relacionales SQL (Structured Query Language).
- Las consultas a la base de datos suelen ser muy rápidas.
- Existe disponibilidad para cualquier sistema operativo, entre los que sobresalen Linux y Unix.
- Establece una conectividad segura.
- Soporta una gran cantidad de datos y de índices por tabla.
- Ejecuta la búsqueda e indexación de campos de texto.

Las ventajas encontradas en el lenguaje de programación PHP son las siguientes:

1. PHP es un lenguaje creado para desarrollar páginas web dinámicas.
2. Es software libre, es decir, cualquiera puede descargar y utilizar el programa de forma gratuita.
3. La sintaxis de PHP es muy fácil y sencilla, además de que puede combinarse con código HTML y Javascript.
4. Es un lenguaje multiplataforma.
5. Tiene capacidad de conexión con la mayoría de los manejadores de bases de datos, entre los que destacan: MySQL, PostgreSQL, Oracle, DB2 y Microsoft SQL Server.

<sup>3</sup> DBA Support [en línea]. <[http://www.dbasupport.com.mx/index.php?option=com\\_content&view=article&id=120%3Acarter%3%ADstic+de+MySQL&catid=12%3Aadministracion&Itemid=7](http://www.dbasupport.com.mx/index.php?option=com_content&view=article&id=120%3Acarter%3%ADstic+de+MySQL&catid=12%3Aadministracion&Itemid=7)> [Consulta: 21 septiembre 2010].

### **Diseño de la interfaz web del Portal de Indicadores Bibliométricos** BIBLAT

El diseño web es el conjunto de actividades que se realizan desde la conceptualización de un sitio web hasta su implementación; no sólo se refiere a tareas de diseño gráfico, sino también a elementos relacionados con el uso de logotipos, el sistema de navegación, las áreas de interacción y de contenidos.

Como se puede apreciar en las siguientes imágenes, el diseño de este portal es sencillo e intuitivo, de tal forma que el usuario puede navegar fácilmente sin tener la sensación de no saber donde se ubica.



Fuente: Portal de Indicadores Bibliométricos BIBLAT

Este portal ha sido desarrollado pensando no sólo en ofrecer calidad en los contenidos, sino también que dichos contenidos sean fácilmente accesibles; por lo tanto, el diseño web está basado en un eficiente sistema de navegación que hace posible la accesibilidad a la información: una sección de contenidos, en la cual el usuario puede visualizar la información e interactuar con ella; una sección de encabezado, de tal forma que el usuario siempre sabe donde está ubicado, y la sección de pie de página, en la cual se hace referencia a los derechos sobre el sitio web y su contenido; todo esto conforma una interfaz sencilla e intuitiva para cualquier usuario.

### **Mantenimiento y actualización de bases de datos**

Dentro de las tareas habituales de mantenimiento de las bases de datos se encuentran aquellas destinadas al control y respaldo de las mismas, como son: control de integridad, chequeo de consistencia y respaldos de seguridad, así como las tareas cuyo objetivo son mantener el *performance* de las bases de datos, como es la reorganización de índices. La frecuencia para realizar estas tareas depende básicamente de tres factores: la cantidad de datos almacenados en las bases, la periodicidad de actualización y el tiempo disponible para ejecutarlas.

En cuanto a la actualización de las bases, ésta es una tarea que se realiza diariamente en el Departamento de Bibliografía Latinoamericana<sup>4</sup> a fin de ofrecer información actualizada a través del web OPAC de CLASE y PERIÓDICA; estas tareas de mantenimiento y actualización deben realizarse directamente en ambas bases de datos dentro del sistema ALEPH 500v.16, para posteriormente hacer las actualizaciones correspondientes en el portal de indicadores bibliométricos desde el manejador de bases de datos MySQL.

<sup>4</sup> Subdirección de Servicios de Información Especializada. Dirección General de Bibliotecas de la UNAM.

## Conclusiones

BIBLAT es una aportación dirigida para un núcleo de usuarios muy especializado, pero su importancia estratégica reside en que da cauce y visibilidad ya no sólo a las revistas y artículos indizados en CLASE y PERIÓDICA, sino a los sujetos que están detrás de la producción científica manifiesta en esas revistas y artículos: los autores y las instituciones que los respaldan, así como los países que sostienen a estas instituciones. Dado que esta información ha adquirido una relevancia cada vez mayor por la creciente tendencia hacia la medición y evaluación del desempeño académico, BIBLAT adquiere un carácter complementario respecto a los criterios y estándares internacionales, pero al mismo tiempo adquiere una pertinencia obligatoria por tratarse de un instrumento que permite compensar los sesgos de los índices internacionales, ya que compila y hace constar la producción de individuos, instituciones y temáticas, que de otro modo pasaría injustamente inadvertida. Ciertamente, BIBLAT requiere de la retroalimentación de esos mismos especialistas hacia los cuales está dirigido, puesto que ello permitirá dar los pasos siguientes hacia la afinación y una mayor sistematización de los reportes requeridos, así como su visualización gráfica de acuerdo con la metodología adecuada. Mientras tanto, la contribución de BIBLAT puede sintetizarse en los puntos que a continuación se enumeran:

- BIBLAT proporciona indicadores de producción científica *alternativos y complementarios* a los indicadores bibliométricos generados por los grandes sistemas de información internacionales.
- La mayoría de los estudios bibliométricos que se han utilizado para analizar la estructura, el desarrollo y las tendencias de la investigación latinoamericana han utilizado como fuente de información las bases de datos de *Web of Science* (ISI-Thomson) y *Scopus* (Elsevier), las cuales no son lo suficientemente representativas de la producción latinoamericana puesto que registran un número muy limitado de revistas de la región.
- La especialización regional de BIBLAT favorece particularmente a las áreas del conocimiento con mayor vocación local como las *Ciencias Sociales, Medicina y Agronomía*.
- BIBLAT ha sido creado aprovechando la sinergia entre dos plataformas: el sistema *ALEPH 500 + software libre*.
- Mediante la plataforma *ALEPH* se hace posible el asentamiento, sistematización y normalización de la información de acuerdo con los criterios de indización de la información y el empleo del formato MARC.
- El *SOFTWARE LIBRE* ha creado la posibilidad de incrementar la capacidad de recuperación y combinación de campos, con objeto de obtener indicadores bibliométricos de la producción científica reflejada en las publicaciones periódicas.
- BIBLAT también ha facilitado la integración de la información a los metabuscadores de Internet como *Google*, potenciando con ello su visibilidad y acceso. ☞



## Obras consultadas

CASTORINA, Jorge A., LÓPEZ LIOTTI, Pablo. *Micro Isis a su alcance: guía rápida*. Buenos Aires: Novedades educativas, 1997. 123 p.

*Citas Latinoamericanas en Ciencias Sociales y Humanidades CLASE* [en línea]. <<http://clase.unam.mx>> [Consulta: 20 septiembre 2010].

*DBA Support* [en línea]. <[http://www.dbasupport.com.mx/index.php?option=com\\_content&view=article&id=120%3ACaracter%C3%ADsticas+de+MySQL&catid=12%3AAdministracion&Itemid=7](http://www.dbasupport.com.mx/index.php?option=com_content&view=article&id=120%3ACaracter%C3%ADsticas+de+MySQL&catid=12%3AAdministracion&Itemid=7)> [Consulta: 21 septiembre 2010].

*Estudio comparativo de las universidades mexicanas* [en línea]. Universidad Nacional Autónoma de México, Dirección General de Evaluación Institucional. <<http://www.ecum.unam.mx/>> [Consulta: 27 agosto 2010].

GIBBS, W W. Lost science in the Third World. *Science*, 1995, vol. 2, no. 273, p. 76-83.

*Google libros* [en línea]. <<http://books.google.com.mx>> [Consulta: 20 septiembre 2010].

*Guía para el Desarrollo de Sitios Web* [en línea]: *versión 2.0*. Gobierno de Chile. <<http://www.guiaweb.gob.cl/guia-v2/>> [Consulta: 21 septiembre 2010].

*Índice de Revistas Latinoamericanas en Ciencias PERIÓDICA* [en línea]. <<http://periodica.unam.mx>> [Consulta: 20 septiembre 2010].

*ISI Web of Knowledge* [en línea]. <<http://isiknowledge.com>> [Consulta: 27 agosto 2010].

*Ranking Iberoamericano SIR 2010* [en línea]. Scimago Institutions Rankings. <[http://www.scimagoir.com/pdf/ranking\\_iberamericano\\_2010.pdf](http://www.scimagoir.com/pdf/ranking_iberamericano_2010.pdf)> [Consulta: 27 agosto 2010].

*RICYT* [en línea]: *Red de Indicadores en Ciencia y Tecnología Iberoamericana e Interamericana*. <<http://www.ricyt.org>> [Consulta: 6 septiembre 2010].

*scimago Journal & Country Rank* [en línea]. <<http://www.scimagojr.com>> [Consulta: 27 agosto 2010].