

IIº ENCUENTRO IBEROAMERICANO DE EDITORES CIENTIFICOS -EIDEC 2010-

Biblioteca Nacional. Buenos Aires, 11 y 12 de noviembre de 2010.

Indicadores cibernéticos: Posicionando la presencia web de las instituciones académicas y de investigación en Latinoamérica

Isidro F. Aguillo

The Cybermetrics Lab

IPP-CSIC

Albasanz, 26-28. 28037 Madrid. Spain

Isidro.aguillo@cchs.csic.es

Resumen: La cibermetría o webometría es una joven disciplina heredera directa de las ciencias cuantitativas dedicadas a la descripción y evaluación de la actividad y resultados de la investigación científica. La bibliometría ha jugado un papel relevante no solo en la comprensión de los procesos de comunicación científica sino incluso en la mejora de los mismos, promoviendo las publicaciones de calidad. Sin embargo, la irrupción de la web, su indudable impacto en la forma de comunicar, su mayor flexibilidad y audiencias gigantescas hicieron necesario un cambio sustancial en las herramientas de evaluación que había que utilizar. La publicación electrónica ya no solo utiliza mecanismos formales como las revistas, ya no describe únicamente resultados sin atender a los datos y los procesos, sino que tiene en cuenta todas las actividades del investigador y docente, ilustrando sus circunstancias, dedicación, contactos, referencias formativas e influencias doctrinales y metodológicas.

Los indicadores web son más complejos, posiblemente más ruidosos y aun es necesaria mucha investigación teórica y práctica en esta disciplina, pero pueden complementar e incluso sustituir los bibliométricos en ciertas misiones y tareas. Son útiles para describir instituciones heterogéneas como las universitarias, disponen de poderosas herramientas para evaluar el impacto tales como el análisis de enlaces y, por primera vez, ofrecen una posibilidad real de estudiar el consumo de información, analizando las visitas a las páginas web académicas y de investigación. Una clasificación de los principales indicadores y las herramientas que se utilizan para su construcción formará el núcleo de la presentación.

La parte empírica deriva de la exitosa experiencia del ranking Web de Universidades, centrando el análisis en las instituciones de la región y destacando el posible impacto de la adopción de políticas de acceso abierto y la implantación de repositorios en las universidades de la región.

Palabras clave: cibermetría, indicadores web, universidades, evaluación, ranking

La cibermetría o webometría es una joven disciplina heredera directa de las ciencias cuantitativas dedicadas a la descripción y evaluación de la actividad y resultados de la investigación científica. La bibliometría ha jugado un papel relevante no solo en la comprensión de los procesos de comunicación científica sino incluso en la mejora de los mismos, promoviendo las publicaciones de calidad. Sin embargo, la irrupción de la Web, su indudable impacto en la forma de comunicar, su mayor flexibilidad y audiencias gigantescas hicieron necesario un cambio sustancial en las herramientas de evaluación que había que utilizar. La publicación electrónica ya no solo utiliza mecanismos formales como las revistas, ya no describe únicamente resultados sin atender a los datos y los procesos, sino que tiene en cuenta todas las actividades del investigador y docente, ilustrando sus circunstancias, dedicación, contactos, referencias formativas e influencias doctrinales y metodológicas.

Los indicadores web son más complejos, posiblemente también más ruidosos y aun es necesaria mucha investigación teórica y práctica en esta disciplina, pero pueden complementar e incluso sustituir los bibliométricos en ciertas misiones y tareas. Son útiles para describir instituciones complejas como las universitarias, disponen de poderosas herramientas para evaluar el impacto tales como el análisis de enlaces y, por primera vez, ofrecen una posibilidad real de estudiar el consumo de información, analizando las visitas a las páginas web académicas y de investigación. De hecho ya existen clasificaciones de los principales indicadores y las herramientas que se utilizan para su construcción.

Tabla 1. Clasificación de indicadores web y equivalentes bibliométricos

Categoría	Cibermetría	Bibliometría	Herramientas web
ACTIVIDAD	Presencia Web	Publicaciones	Páginas web en motores de búsqueda Motores académicos
IMPACTO	Visibilidad Web	Citas	Análisis de enlaces Análisis de menciones
CONSUMO	Visitas/volcados	Circulación	Análisis de tráfico Análisis de ficheros log

Una de las aplicaciones de los indicadores cibernéticos que ha tenido mayor impacto es el ranking web de universidades, que analiza y clasifica las instituciones de educación superior de todo el mundo y en el que entre otros factores evalúa el impacto de la adopción de políticas de acceso abierto y la implantación de repositorios en las universidades.

Antecedentes del Ranking

A mediados de los noventa el Laboratorio de Cibermetría del CSIC (Consejo Superior de Investigaciones Científicas de España) inició el desarrollo de indicadores web cuantitativos utilizando técnicas similares a las bibliométricas y cienciométricas. La investigación se centró en las instituciones académicas y de investigación aunque los primeros resultados no fueron muy prometedores dado que muchas de las universidades apenas tenían presencia destacada en la web en ese momento. Además el principal indicador usado entonces, el factor de impacto web de ingwersen presentaba notables sesgos y carencias técnicas. La recolección sistemática de datos se inició en 1999 utilizando robots, pero dado los elevados requerimientos de recursos humanos y de cómputo, finalmente se optó por utilizar las bases de datos de los principales motores de búsqueda comerciales.

En 2003 aparece el llamado Ranking de Shanghái, que utiliza un indicador compuesto que sirve de modelo para su aplicación al entorno web. En 2004 aparece la primera versión del Ranking Web (www.webometrics.info), o Ranking Webometrics con el fin de promover la difusión en abierto del conocimiento científico, las llamadas iniciativas Open Access. El ranking es el resultado de una serie de trabajos de investigación y por tanto ha ido evolucionando con el paso del tiempo, adecuando la metodología a los resultados obtenidos con el obvio sacrificio de la estabilidad de las posiciones entre ediciones consecutivas. Sin embargo, las bases fundamentales ya están establecidas a partir de las ediciones de 2006.

Hipótesis central

La Web es indiscutiblemente la más importante fuente de información académica en la actualidad. Es universal, fácil de acceder, económica para publicar y puede alcanzar audiencias gigantescas y muy diversas en cualquier lugar del mundo. Los resultados publicados obtienen un impacto inmediato y no hace falta que se distribuyan a través de canales formales, pudiendo incorporar material adicional imposible de suministrar en papel bien por razones técnicas (ficheros multimedia) o económicas (acceso a los datos originales).

Pero su utilidad va mucho más allá de la comunicación científica ya que es el escaparate perfecto para reflejar la estructura, actividades y resultados de cualquier

organización, incluyendo explícitamente las de carácter académico. Nuestra hipótesis es que si la presencia web refleja el desempeño global de una Universidad, la calidad docente e investigadora de sus departamentos, la extensión de sus servicios y el prestigio de la institución, entonces un ranking basado en indicadores web puede clasificar de manera fidedigna las diferentes instituciones.

Se trata de una herramienta que puede ayudar al diseño de estrategias a medio y largo plazo, pero también a la adopción de medidas concretas inmediatas tales como las políticas de acceso abierto que tienen un indudable impacto en la visibilidad web de una universidad.

Si la hipótesis es válida, entonces los indicadores web pueden ser utilizados para la descripción y evaluación de universidades y organizaciones similares, proporcionando herramientas adicionales al arsenal de indicadores ya disponibles. Incluso si las políticas web no son prioritarias aún para ciertas instituciones, el ranking está creando un círculo virtuoso empujándolas a mejorar su presencia en la web.

De forma sucinta el Ranking se ha construido de acuerdo a los siguientes principios:

- El Ranking Web es un ranking, una clasificación ordenada de universidades según un único indicador compuesto. El indicador combina distintas variables seleccionadas para proporcionar una indicación, simple y sintética de las prestaciones globales de una universidad.
- El Ranking Web es único. Se ha elegido una única combinación de pesos y variables que refleja en opinión de los autores la mejor síntesis. El ranking es tamaño dependiente y no pretende reflejar prestaciones relativas (eficiencia). La contribución relativa de cada variable debe ser objeto de análisis individual y no pretende ser es objeto del ranking actual.
- El Ranking Web es mundial. El objetivo es analizar todas las universidades del mundo (más de 20,000), especialmente las de los países emergentes y en vías de desarrollo.
- El Ranking Web tiene en consideración todas las misiones universitarias. Los resultados de investigación es uno de los principales indicadores, especialmente para la medida de la excelencia de las llamadas universidades “World-class”, pero el resto de las misiones tienen que ser también tenidas en cuenta. La docencia (especialmente a distancia) y la llamada tercera misión (transferencia) pueden ser evaluadas indirectamente a través de la presencia web. El análisis de enlaces se ha demostrado muy potente para descubrir este tipo de actividades.

No se han usado técnicas cualitativas que quizás sean apropiadas para poblaciones pequeñas pero que resultan inviables o poco fidedignas para describir mediante resultados de encuestas los escenarios globales propuestos.

Una de las ventajas de los indicadores web es que están involucradas cifras muy elevadas: cientos de miles de páginas web y millones de enlaces hipertextuales. Ello significa que aunque las motivaciones para enlazar son muy variadas y a veces espúreas o irrelevantes el ruido producido no afecta los principales patrones estadísticos que se ven reforzados por la ley de los grandes números. Es más, en dicha diversidad es posible descubrir actores ajenos a los círculos académicos que pueden ser muy relevantes en la organización del sistema, y que han sido habitualmente ignorados en los trabajos bibliométricos

Por último cabe señalar la capacidad de segregación (en el ranking) de estos indicadores frente a la que producen por ejemplo el número de premios Nobel que apenas puede clasificar unas pocas docenas de universidades.

Construyendo el Ranking

Se calcula que pueden existir entre 22.000 y 24.000 instituciones de educación superior, de las que tres cuartas partes pueden clasificarse como universidades. Nuestro Directorio incluye alrededor de 20.000 entradas de instituciones con presencia web y dominio propio, con cerca de un millar con más de un dominio principal (por ejemplo, manteniendo el dominio antiguo tras cambiar a uno nuevo) o una sección con dominio diferente.

Dado el gran tamaño de la población y la necesidad de obtener varias variables, la captura de datos se realiza de forma indirecta extrayendo la información de las bases de datos de los principales motores comerciales de búsqueda. Google, Bing, Yahoo, Exalead y Google Académico ofrecen herramientas y operadores para la captura automática de indicadores web. Se ha desarrollado una estrategia para minimizar los problemas de cobertura y el comportamiento irregular de algunos de los motores, que incluye la colecta repetida en un corto periodo de tiempo.

El indicador compuesto (WR) se ha diseñado siguiendo el modelo del factor de impacto, que de forma general divide citas (visibilidad) entre artículos (actividad). En la Web la visibilidad corresponde a los enlaces externos recibidos mientras que la actividad se refiere a las páginas web publicadas. Puesto que las variables siguen leyes de potencia y son más diversas en su composición resulta más adecuado utilizar un polinomio en vez de un

índice pero manteniendo la relación 1:1 entre ambas variables: visibilidad (50%) y actividad (50%).

Un examen de los componentes de la mediada de actividad determinó la utilidad de segregar dicha variable en tres para reflejar específicamente la actividad de comunicación científica tanto informal como formal. En este último caso la libre disponibilidad de una base de datos bibliográficos con citas (Google Académico) fue clave. En la actualidad el indicador se calcula como se indica a continuación:

- **Actividad:** Se consideran tres componentes, tamaño medido en número de páginas web (20%), número de documentos contando los ficheros en determinados formatos tales como pdf, doc, ppt o ps (15%) y número de artículos (15%) utilizando el ya citado Google académico.
- **Visibilidad.** El 50% restante se calcula a partir de los enlaces externos recibidos, especialmente si vienen de otros dominios académicos.

Hasta 2009 los pesos se aplicaban a los valores ordinales, pero desde esa fecha se hace de una forma más lógica sobre los datos normalizados para cada variable.

Resultados

En general se observa una apreciable correlación entre los resultados proporcionados por el Ranking Web y los de otros rankings mundiales, a pesar de las diferencias en criterios y métodos.

Hay varias causas que explican las discrepancias encontradas:

- Criterios de selección e inclusión. Algunos ranking tienen coberturas restringidas o abiertamente excluyen ciertas universidades
- Todos los rankings cometen errores puntuales y están basados en Fuentes no completamente fidedignas con algunos sesgos muy obvios
- El cambio de dominio web, la fusión o partición de una institución y su reflejo en la web tienen efectos catastróficos sobre su posición
- Iniciativas concretas como grandes repositorios, portales con amplios índices, directorios o bases de datos pueden hacer destacables universidades cuyas demás prestaciones no son tan notables.

La Tabla 2 proporciona un resumen de los resultados distribuidos por regiones, señalando específicamente los países latinoamericanos, que sólo sitúan 3 universidades entre las 200 primeras y 9 entre las Top 500, fundamentalmente mejicanas y brasileñas).

Tabla 2. Cifras principales del Ranking Web, Julio 2010

Región/País	Top 100	Top 201	Top 500	Top 1000	TOTAL
NORTE AMÉRICA	79	123	212	398	3485
EUROPA	16	60	222	415	4975
ASIA	3	7	38	104	6142
OCEANIA	1	6	15	35	149
LATINOAMÉRICA	1	3	9	39	3486
México	1	1	2	5	907
Brasil		1	5	17	1494
Chile		1	1	3	81
Argentina			1	3	105
Colombia				3	282
Perú				2	93
Venezuela				2	55
Costa Rica				1	60
Puerto Rico				1	40
Ecuador				1	70
MUNDO ARABE	0	2	3	4	569
AFRICA	0	0	1	5	355
TOTAL					19161

También es interesante observar la brecha abierta entre las universidades norteamericanas y las europeas, especialmente cuando nos centramos en las instituciones líderes. Aunque se trata de un hecho preocupante hay que considerar diversos factores, no siendo el menor la existencia de grandes consejos de investigación en Europa (CNRS en Francia, CNR en Italia, Max Planck en Alemania o CSIC en España) que aunque a menudo colaboran con las universidades son independientes de estas y son tremendamente productivos.

Las universidades latinoamericanas aparecen en el Ranking Web mejor posicionadas que en otros rankings similares. Frente al peso de la investigación en las llamadas “World-class”, se trata de instituciones que vertebran el país y que por tanto de gran cumplen otras importantes misiones. El liderazgo corresponde a universidades de gran tamaño, con una gran tradición histórica y que en algunos casos han apostado por invertir en enseñanza no presencial. Sin embargo falta aún mucho por hacer incluyendo el desarrollo de políticas de acceso abierto más beligerantes, el relanzamiento de la misión investigadora o alcanzar masas críticas en algunos países con un sistema muy fragmentado.

Otros Rankings

El Laboratorio de Cibermetría produce además del Ranking Web de Universidades (www.webometrics.info), el de Centros de Investigación (research.webometrics.info), Hospitales (hospitals.webometrics.info), repositorios (repositories.webometrics.info) y Escuelas de negocios (business-schools.webometrics.info).

Bibliografía

Aguillo, I.F., Bar-Ilan, J., Levene, M., Ortega, J.L. (2010). Comparing university rankings. *Scientometrics*, 85 (1): 243-256.

Aguillo, I. (2009). Measuring the institution's footprint in the web. *Library Hi Tech*, 27 (4), pp. 540-556.

Ortega, J.L., Aguillo, I.F. (2009). Mapping world-class universities on the web. *Information Processing and Management*, 45 (2), pp. 272-279.

Aguillo, I.F.; Ortega, J. L. & Fernández, M. (2008). Webometric Ranking of World Universities: Introduction, Methodology, and Future Developments. *Higher Education in Europe*, 33(2/3): 234-244.

Aguillo, I. F.; Granadino, B.; Ortega, J. L.; Prieto, J. A. (2006). Scientific research activity and communication measured with cybermetric indicators. *Journal of the American Society of Information Science and Technology*, 57(10): 1296 – 1302.