

CONTRIBUCIÓN DEL ACCESO ABIERTO A LA VISIBILIDAD DE LA LITERATURA CIENTÍFICA EN UNA INSTITUCIÓN DE EDUCACIÓN SUPERIOR

C. G. Pavão¹, J. S. B. da Costa¹, Z. Horowitz¹, M. K. Ferreira¹, S. E. Caregnato²

¹Centro de Processamento de Dados, UFRGS - Brasil, ²Faculdade de Biblioteconomia e Comunicação, UFRGS - Brasil

comissao@cpd.ufrgs.br, sonia.caregnato@ufrgs.br

Resumen

La visibilidad es una cuestión crucial, no sólo para los investigadores, sino también para sus instituciones, teniendo en cuenta que los recursos invertidos en investigaciones científicas sólo se justifican si sus resultados son ampliamente accesibles. Los repositorios institucionales pueden constituir un instrumento de reordenamiento de los criterios de visibilidad de la literatura científica, del investigador y de las instituciones, porque va más allá de los mecanismos de difusión hasta ahora utilizados. Este trabajo buscó constatar la visibilidad ofrecida por un repositorio a la producción científica de una institución de educación superior mediante el análisis de los accesos y descargas. Son indicados los desarrollos incorporados en las estadísticas de Lume - Repositorio Digital de la Universidad Federal de Rio Grande do Sul, con el fin de facilitar y promover su uso por investigadores y gestores. En el texto se exponen los mecanismos disponibles en Lume que aportan para su consolidación como instrumento de gestión de las investigaciones, por permitir análisis más refinados respecto de la visibilidad de las investigaciones realizadas en la institución. Se concluye que los repositorios institucionales pueden contribuir con los nuevos procesos de comunicación científica.

Palabras Clave: Acceso abierto, Producción científica, Visibilidad, Repositorios institucionales, Instituciones de educación superior.

Abstract

Open access has been increasing significantly the visibility of current research. Thus, researchers have pressured publishers to remove barriers for this access. Visibility has become a crucial issue, not only for researchers, but also for their institutions, since spending could be only justified if its results are widely accessible. However, this new vision on the validity of scientific publication challenges established models of scientific communication, especially regarding the traditional roles for its actors and formats of scientific publication. From this perspective, the institutional repositories may constitute an instrument of rearrangement, enhancing visibility of scientific literature, researchers and institutions, going beyond the

publicity mechanisms used to date. This study tries to demonstrate the visibility provided by a repository of scientific production of a higher education institution, through access and download analysis. The developments incorporated by the statistics of Lume - Digital Repository of Federal University of Rio Grande do Sul, Brazil, are presented with the purpose of facilitating and promoting its use by researchers and managers. It shows the mechanisms available in Lume that contribute to its consolidation as a tool for management of research, by allowing more precise analysis regarding visibility of research available at the institution. It concludes that institutional repositories can help a lot in the new processes of scientific communication.

Keywords: open access; scientific production; visibility; institutional repositories; higher education institutions.

INTRODUCCIÓN

Muchas instituciones brasileñas y extranjeras han creado sus repositorios digitales, con el fin de asegurar el almacenamiento permanente y fiable de la producción científica generada a partir de sus actividades de enseñanza e investigación, para permitir la consulta rápida y eficiente, así como aumentar la visibilidad y el uso de lo que se produce en su ámbito.

En este trabajo se presentan los recursos disponibles en Lume - Repositorio Digital de la Universidad Federal de Rio Grande do Sul, como herramienta para la verificación de la visibilidad de los resultados de las investigaciones desarrolladas por la Institución. El objetivo es contribuir con las demás instituciones de educación superior y de investigación en el desarrollo de mecanismos para la recolección de los datos estadísticos que permitan la evaluación y comparación de la visibilidad de la investigación en las instituciones y entre ellas. La visibilidad puede ser verificada por el análisis de los datos de acceso y descarga que permiten mostrar el interés por la investigación y la aceptación de las mismas, tanto en el ámbito nacional como internacional.

En este sentido, se describe la implementación e implantación de un módulo responsable por la captura de datos y que presenta, de una forma eficiente, las estadísticas de acceso y de descargas de los objetos digitales de Lume. Después de las adaptaciones en el módulo, las estadísticas pasaron a presentarse de diversas formas: agrupadas por comunidad o colección, por autor o palabra clave, o a partir de un conjunto de informaciones resultantes de una búsqueda.

Lume ha sido desarrollado utilizando DSpace - *Institutional Repository System* [1], un programa abierto que ofrece un portal web combinado con un banco de datos para archivo y consulta de las colecciones, con soporte de los diversos formatos de archivo y organización en comunidades que se subdividen en subcomunidades y colecciones.

Originalmente, DSpace tiene un módulo de estadísticas que muestra información relativa a todo el repositorio, sin mucha sofisticación; proporciona datos sobre los totales generales, tales como el total de

accesos y descargas, el total de ítems que forman el acervo, los ítems más visualizados, los usuarios más frecuentes, además de la información técnica.

El módulo de estadísticas elaborado para Lume tiene una manera más completa y eficaz de presentar la información: se visualizan gráficos de los datos estadísticos sobre un ítem, colección, comunidad, autor, palabra clave o de un grupo de ítems resultantes de las búsquedas. También están disponibles opciones para limitar el período de consulta de las estadísticas. Los registros de acceso se guardan en una tabla de *log* en la base de datos y los registros estadísticos son actualizados periódicamente por medio de la aplicación de algunos *scripts* que examinan los registros de *log* de la tabla y generan los datos mencionados. Además, el módulo tiene una presentación más completa, con el uso de la interfaz XMLUI, donde son exhibidos los gráficos de datos estadísticos para cada ítem, colección o comunidad, y están disponibles las opciones para limitar el período de consulta de las estadísticas.

En la primera parte de este texto se presentan algunas consideraciones sobre comunicación científica; en la siguiente sección se tratará el tema de la visibilidad; luego, serán abordados los aspectos técnicos para la implementación del módulo de estadísticas; la cuarta sección describe la implementación de la presentación gráfica de las estadísticas; y finalmente, se presentan las conclusiones.

COMUNICACIÓN CIENTÍFICA

El ritmo acelerado con que el cambio científico y tecnológico se produce actualmente ha hecho que muchos de los resultados de los estudios publicados en canales formales de comunicación estén obsoletos cuando son publicados. Ya en el año 2000, Campello [2] aseguraba que esto ha privilegiado el surgimiento de nuevos medios de difusión del conocimiento científico, principalmente gracias a los medios de comunicación electrónicos de divulgación.

La aceleración de la producción de conocimiento está preocupando también a las instituciones y agencias de fomento, puesto que difundir ampliamente las investigaciones evitan gastos innecesarios con la investigación repetitiva y permite el uso de resultados para acelerar los avances de la ciencia. Las universidades, los institutos y los centros de investigación están cada vez más preocupados con la difusión de la ciencia y, en consecuencia, se han comprometido a implementar mecanismos para facilitar la difusión de lo que está siendo desarrollado por sus investigadores. Entre estos mecanismos está el fomento a la creación de repositorios institucionales (RI).

Los repositorios han surgido en la década de los 90, cuando las revistas científicas son el principal medio de divulgación de la producción científica de los investigadores y las universidades. Pueden considerarse, por lo tanto, como alternativa o complemento a los sistemas tradicionales de comunicación científica, que se encontraba de alguna manera restricta y elitista, debido a los altos costos de publicación y de adquisición por parte de las bibliotecas y los investigadores [3].

El aumento de la mercantilización de la información por las grandes editoras, para las cuales los investigadores entregan de forma gratuita los resultados de su trabajo, financiado con fondos públicos de las universidades o de las agencias de fomento, para que regresen a las universidades, vendidas a un precio bastante alto, guió el movimiento en favor del acceso abierto a la información [4].

Con la difusión, las producciones de los investigadores pueden ser evaluadas por importantes indicadores que consideren la cantidad y la calidad de la información que el investigador comunica.

En el caso de los investigadores académicos, una manera de evaluar la productividad es medir la cantidad de artículos de revistas que publican y una forma de evaluar la calidad de la producción científica consiste en determinar el grado de interés que los otros tienen por la investigación. Esto puede hacerse mediante el control del número de citas que un trabajo recibe en la bibliografía ulterior. La importancia de una publicación se determina por el hecho de ser altamente citada durante un largo período de tiempo. Para que esto ocurra deben estar asociadas algunas características, como la novedad de la investigación, un campo de cambio rápido, la absorción de conceptos por la conciencia general más o menos rápida y desarrollos importantes [5]. Podemos añadir la tendencia de aumento en las citas de publicaciones de acceso libre.

Las revistas científicas han sido el principal medio utilizado por los autores para la divulgación de su producción. Esta preferencia está garantizada por el prestigio que históricamente este tipo de publicación proporciona a la comunidad científica. La prioridad por este tipo de publicación depende, en gran parte, de la visibilidad alcanzada a partir de la publicación de un artículo en una revista. El grado de notoriedad al cual los investigadores tratan de llegar repercute en los canales de comunicación, las instituciones y los tipos de producción científica.

El modelo de archivos abiertos, en desarrollo en diversos países, tiene como objetivo hacer más eficiente la difusión de la información académica y aumentar la visibilidad de la investigación desarrollada en las universidades.

Los notables avances tecnológicos en esta área y el creciente uso de RI para gestionar colecciones digitales han hecho crecer también la necesidad de establecer mecanismos para controlar la visibilidad lograda por los documentos depositados, con el fin de utilizar los datos en la gestión académica, de la investigación, de los recursos invertidos y en la evaluación institucional.

VISIBILIDAD

A partir de la perspectiva anterior, el investigador podrá tener en cuenta otras formas de visibilidad, frente a la mayor accesibilidad y el impacto de lo que produce. El concepto de visibilidad (*visibility*) puede ser interpretado como un tipo específico de capital social. La visibilidad [...] expresa el valor diferencial, distintivo, de esta especie particular de capital social: acumular capital es hacer un “nombre”, un nombre propio, un nombre conocido y reconocido, marca que distingue inmediatamente a su portador,

removiéndolo para una forma visible de lo indeterminado, inadvertido, oscuro, en el que se pierden los hombres comunes [6].

La visibilidad se está convirtiendo en una cuestión crucial, no sólo para los investigadores, sino también para las instituciones. El acceso abierto aumenta considerablemente la visibilidad de la investigación y los investigadores han ejercido presión sobre los editores de sus trabajos para que eliminen las barreras de acceso, porque entienden que los recursos asignados a la investigación sólo se justifican si sus resultados son ampliamente accesibles. Sin embargo, esta nueva visión sobre la legitimidad de la publicación científica desafía los modelos ya establecidos. Por lo tanto, debemos encarar el desafío e intentar cambiar el modelo del proceso de comunicación científica vigente, tanto en términos de los roles tradicionales de sus actores como al respecto de los formatos de publicación científica [7].

Por esta perspectiva, los RI pueden ser el instrumento de reordenación de los criterios de visibilidad de las publicaciones científicas, del investigador y de las instituciones, ya que ahora se pasará a producir para un universo que extrapola las dimensiones de los mecanismos de difusión hasta ahora utilizados. La visibilidad que los RI proporcionan puede atraer más atención hacia los investigadores que están depositando su producción científica en ellos, en relación con los colegas que no suelen utilizar esta herramienta como medio para divulgar sus investigaciones.

De esta manera, los repositorios institucionales tienen el poder de interferir, principalmente, en dos aspectos estratégicos [8]:

- Contribuir con el aumento de la visibilidad y “valor” público de las instituciones, sirviendo como un indicador tangible de su calidad.
- Contribuir para la reforma del sistema de comunicación científica, ampliar el acceso a los resultados de investigaciones académicas y recuperar el control sobre la publicación científica.

La visibilidad científica se define como el grado de exposición y la evidencia de un investigador frente a la comunidad científica. Una posición de alta visibilidad es aquella en la cual la obra y las ideas del investigador son fácilmente accesibles. La falta de visibilidad sería la situación en la cual las obras y las ideas de un investigador no son conocidas ni tienen la oportunidad de serlo, por no estar disponibles en canales de fácil acceso. Las posibilidades de un autor para incrementar su visibilidad dependen, entonces, de su mayor exposición a la comunidad mundial [9].

La visibilidad internacional es uno de los varios indicadores utilizados por los sistemas de evaluación para determinar la posición de prestigio de un investigador. [...] las comunidades científicas y cada autor miembro, independientemente de la situación de desarrollo científico de su país, tienen las mismas aspiraciones de lograr mayor visibilidad, reconocimiento y citas. Por esta razón, según la tradición profundamente inculcada en cada investigador durante el proceso de su formación, todos sus esfuerzos se orientan a publicar, sobre todo en las revistas más prestigiosas, generalmente internacionales e indexadas. En Brasil, como en muchos otros países, estos esfuerzos se han estimulado y recompensado por las

universidades y los organismos de desarrollo. No siempre, sin embargo, los esfuerzos tienen éxito. Una alternativa muy atractiva, pero sin duda utópica, podría ser ofrecida si todas las universidades del mundo apoyasen los repositorios institucionales de acceso libre; todos tendríamos acceso a todo. [10]. En relación con el sentido de visibilidad internacional, puede decirse que es aquello que se propaga por varias naciones [11].

Las instituciones mantenedoras de los repositorios acogen, entre otras funciones, la función primordial de aumentar la visibilidad de la institución y de sus investigadores, permitiendo y fomentando el acceso a su producción científica, aumentando las posibilidades de que sean leídas, evaluadas y citadas.

Se considera que la visibilidad de las publicaciones de un país en la literatura técnica y científica internacional es un indicador importante de su participación en el escenario científico internacional. La visibilidad científica es, claramente, un indicador importante del desarrollo y reconocimiento internacional de la producción científica de un país. En el caso de los países en desarrollo, el aumento de la visibilidad de sus autores hace posible la citación de los autores por parte de otros colegas de países centrales y eso significa una inversión en la dirección de la influencia esperada dentro del modelo centro-periferia. Países que antes eran apenas consumidores de la ciencia producida por otros, son ahora participantes en la producción científica mundial, dejando de lado el aislamiento científico internacional [12]. A los países alejados de los centros productores, los repositorios trajeron la posibilidad de acceder a lo que se produce en países centrales y la perspectiva de que la producción local tenga una visibilidad mayor y penetración internacional.

Puede haber beneficios sustanciales con la adhesión cada vez mayor del acceso abierto y existen maneras de refinar o incluir mecanismos para comprobar el impacto y la visibilidad de los documentos disponibles en los repositorios; por ejemplo, para elaborar estimaciones más documentadas sobre el aporte de los datos de investigación, publicaciones y otros objetos en acceso abierto a los conocimientos generados; encontrar enlaces directos y mensurables para el acceso, uso y eficacia, para poder estimar el porcentaje de variación apropiadamente y aplicarlo a la variable acceso, descarga y citaciones; encontrar medidas para comparar el uso de los contenidos de acceso, basadas en la suscripción y el acceso abierto para los no-usuarios de investigación (por ejemplo, los usuarios de la industria, los gobiernos, las organizaciones no gubernamentales y la comunidad en general); explorar la ampliación del alcance y el impacto más allá de las fronteras nacionales, con el fin de comprender mejor desde dónde pueden obtenerse los beneficios y la importancia de las iniciativas internacionales de acceso abierto; llevar a cabo el análisis de los posibles efectos del cambio de acceso para explorar el promedio de las tasas de retorno del desarrollo de la ciencia y la tecnología, y el impacto del aumento de la accesibilidad y la eficiencia. [13]

Con el objetivo de proporcionar mecanismos de verificación de la visibilidad de los productos de las investigaciones realizadas en la Universidad Federal de Rio Grande do Sul (UFRGS), se crearon las estadísticas de uso de Lume.

Es importante mencionar que muchos estudios demuestran que los trabajos en acceso abierto, incluidos en los RI, se emplean con mayor frecuencia, tanto en términos de descargas como de citas; sin embargo, este trabajo no contempla el foco de las citas [13]. Por lo tanto, además de los datos estadísticos de los accesos y descargas, extraídos de los RI, es necesario el análisis de citas utilizando herramientas como el *software Publish or perish* o bases de datos que permitan este uso, tales como *Web of Science*, *Scopus*, entre otros.

IMPLEMENTACIÓN TÉCNICA DEL MÓDULO DE ESTADÍSTICAS

Lume tiene como objetivo reunir, conservar, difundir y proporcionar acceso confiable y permanente a los documentos académicos y científicos de la Universidad. Actualmente usa la versión 1.5.2 de DSpace instalado en un servidor Linux. El DSpace está desarrollado en Java y proporciona dos interfaces gráficas: JSPUI y XMLUI. Lume usa la interfaz XMLUI y las siguientes herramientas de código abierto para dar soporte al portal:

- Banco de datos relacional PostgreSQL versión 8.3, donde se almacenan los metadatos (datos sobre datos).
- Servidor de aplicaciones Java Tomcat Servlet Engine versión 6.
- Motor de búsqueda LUCÈNE, para la indización y recuperación de los metadatos y documentos textuales.
- Apache HTTP Server, versión 2.
- Bibliotecas Java (JavaBeans Activation Framework, Java Servlet y Java Mail API).
- Compilador Ant.

Con el fin de garantizar el acceso permanente a los objetos digitales es utilizado el sistema Handle System, gestionado por la CNRI - *Corporation for National Research Initiatives* [15]. Este sistema proporciona un identificador persistente para cada documento digital y garantiza que la URL no tendrá que sufrir alteraciones cuando el sistema migre a un nuevo equipo (hardware), o cuando se produzcan alteraciones en el sistema. Con ello se asegura que, incluso si la dirección del servidor del repositorio digital cambia, los recursos se puedan continuar usando a partir de la URL persistente.

Actualmente, Lume contiene metadatos y objetos digitales provenientes de la captura automática de otros sistemas (*harvesting*); esta práctica permite la integración de los diferentes procesos de depósito que se hacen necesarios en un sistema multidisciplinario:

- El sistema ACERVO FOTOGRÁFICO, desarrollado internamente, contiene las colecciones de fotos de diversos organismos de la universidad.
- El sistema SABI, Sistema de Automatización de Bibliotecas de la UFRGS, utiliza el software Aleph500, que gestiona las actividades y servicios ofrecidos por las bibliotecas a la comunidad de usuarios.

Lume se compone de metadatos e información digital. Los metadatos son información estructurada con semántica estandarizada y se utilizan para representar la información digital en forma de referencias bibliográficas. Sigue el estándar *Dublin Core - Dublin Core Metadata Initiative* [16].

Para asegurar la interoperabilidad con otros repositorios digitales, utiliza el protocolo de captura de metadatos de la iniciativa de archivos abiertos OAI-PMH - *Open Archives Initiative Protocol for Metadata Harversting* [17], lo que permite la transferencia de datos entre repositorios digitales.

Para que las estadísticas sean precisas, los datos deben ser procesados correctamente. El uso de los registros de *logs* requiere mucho cuidado para que los datos recogidos no provoquen distorsiones en los análisis. Pero la forma en que se generan los registros de uso puede variar entre plataformas; por lo tanto, es impracticable describir todos los posibles filtros utilizados para limpiar los datos. En Lume, por ejemplo, los datos sobre los accesos ejecutados por robots de la Web, que hacen búsqueda automática en el repositorio para recoger la información y ponerla a disposición en su propia herramienta de búsqueda, no se registran porque no son accesos verdaderos y aumentan sustancialmente las estadísticas.

Así mismo después de las consideraciones señaladas y de tomar las precauciones necesarias para evitar la distorsión de la información estadística, se considera que los registros de *logs* si son una rica fuente de información fiable para llevar a cabo el análisis de los datos de acceso de una herramienta de búsqueda y recuperación de la información disponible en la web.

En esta sección se describen los incrementos ejecutados en la base de datos y el código fuente de DSpace para posibilitar la captura, el almacenamiento y la presentación de la información relativa a los accesos y descargas realizadas en todo el repositorio o específicamente de un ítem, una comunidad, una subcomunidad, una colección o una determinada búsqueda realizada.

MODIFICACIONES EN EL BANCO DE DATOS

Se han creado cinco nuevas tablas en la base de datos PostgreSQL del DSpace:

- TIPO_ACCESO: Esta tabla registra los tipos de accesos que se utilizarán posteriormente para generar la visualización de las estadísticas: visualización del ítem, de la colección, de la comunidad y descarga de los objetos digitales. En la Tabla 1 se enumeran los tipos de acceso y su descripción.

Tipo_acceso	Descripción_tipo_acceso
1	Visualización del ítem
2	Visualización de la colección
3	Visualización de la comunidad
4	Descargas (<i>Downloads</i>)

Tabla 1. Tipos de acciones que se consideran en las estadísticas

- **ACCESOS:** Esta tabla registra cada acceso ejecutado. Por esta razón, es una tabla grande, que hasta ahora ha alcanzado 19 millones de registros desde el inicio de Lume, pudiendo causar demoras en la respuesta del módulo de estadísticas. Para permitir el uso de estos datos hay dos tablas: **ACCESOS_RESUMEN** y **ACCESOS_AÑO_MES**, tal como se describe a continuación. Estas tablas almacenan la información de la tabla **ACCESO** de forma resumida, agrupan los datos y optimizan el tiempo necesario para mostrar los datos. Ejemplos de datos de la tabla **ACCESO** puede verse en la Tabla II.
- **ACCESO_RESUMEN:** Esta tabla agrupa los registros de la tabla **ACCESOS** con el fin de generar las estadísticas del total de descarga por comunidad/colección y el *ranking* de los ítems con más descargas en cada comunidad de Lume.
- **ACCESOS_AÑO_MES:** En esta tabla se almacenan los datos de la tabla **ACCESOS**, agrupados por año, mes, país, tipo de acceso y *handle*. El objetivo de esta tabla es optimizar el rendimiento del módulo de estadísticas que, inicialmente, consultaba la tabla **ACCESOS**. Esta tabla tiene el 13% del tamaño de la tabla original gracias al agrupamiento de los datos.

seq_acceso	handle	fecha_hora_acceso	ip	nombre_archivo	tipo_acceso	país	año_mes_día
93206	10183/614	2008-01-08 14:20:58.938368	143.54.69.13	000554032.pdf	4	UFRGS	20080108
93207	10183/8992	2008-01-11 07:50:08.131795	143.54.34.79		3	BR	20080111
93208	10183/4914	2008-01-12 10:02:07.471584	200.144.94.132		1	BR	20080112
93195	10183/86	2008-01-08 13:51:16.819651	143.54.69.13		2	BR	20080108

Tabla 2. Ejemplos de registros de acciones que se mantienen en la tabla accesos

- **ACCESOS_LOG_ACTUALIZA:** Esta tabla registra cada actualización de las tablas **ACCESOS_RESUMEN** y **ACCESOS_AÑO_MES**, por medio de una función incorporada en la base de datos.

La Figura 1 muestra el flujo de actualización de las tablas de accesos resumidas en la tabla **ACCESOS**, ejecutada periódicamente.

(Ver figura en la siguiente página)

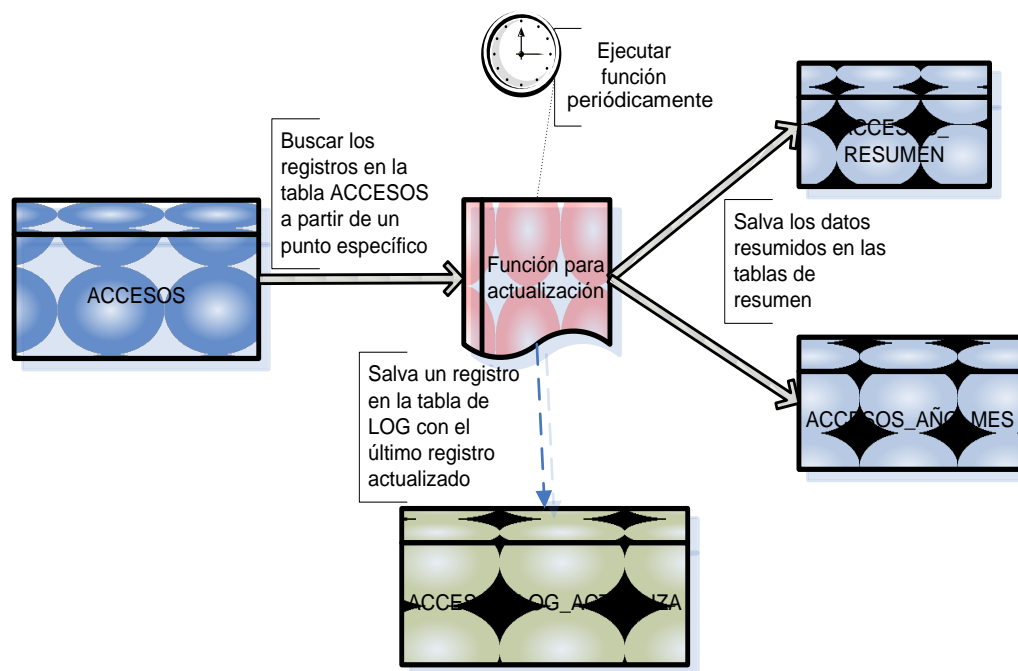


Figura 1. Flujo de la actualización de las tablas de accesos resumidas a partir de la tabla ACCESOS

INCREMENTOS EN EL CÓDIGO FUENTE JAVA

Para cada tipo de acceso ejecutado se creó una clase Java responsable por el registro de los accesos. Estas clases fueron configuradas para ser llamadas a cada carga de una página, con un estándar específico de URL.

Todas estas modificaciones se llevaron a cabo mediante la creación de nuevos módulos que sólo necesitan ser agregados en la jerarquía del directorio original del código fuente de DSpace, pasando a ser consideradas en la compilación efectuada mediante la herramienta Ant.

Esta forma de agregar módulos aprovecha las facilidades proporcionadas por los desarrolladores de DSpace cuando se usa el interfaz XMLUI, donde los incrementos en el código se realizan en una ubicación específica dentro de la jerarquía de directorios del código fuente de DSpace. Cada nuevo módulo agregado se coloca en esta ubicación específica y los ficheros de configuración de la compilación de la herramienta Ant son modificados, con el fin de examinar este nuevo módulo. De esta manera, a cada nueva actualización de la versión de DSpace es sencillo añadir los nuevos módulos, siendo necesario sólo reemplazar los módulos en un lugar determinado, copiando los ficheros de configuración de la herramienta Ant y compilar.

IMPLEMENTACIÓN GRÁFICA PARA LA VISUALIZACIÓN DE LAS ESTADÍSTICAS CAPTURADAS

Las estadísticas son presentadas en el portal de Lume de las siguientes maneras:

- Estadísticas resumidas por comunidad, colección o ítem.
- Estadísticas para ítems de un autor o una palabra clave.
- Estadísticas para ítems de una búsqueda avanzada.
- Estadísticas generales por comunidad o colección.

A. ESTADÍSTICAS RESUMIDAS POR COMUNIDAD, COLECCIÓN O ÍTEM

Estas estadísticas están disponibles para todas las comunidades, colecciones e ítems presentes en el repositorio. Son accesibles por medio de un ícono presente en la página principal de los elementos a los que pertenecen.

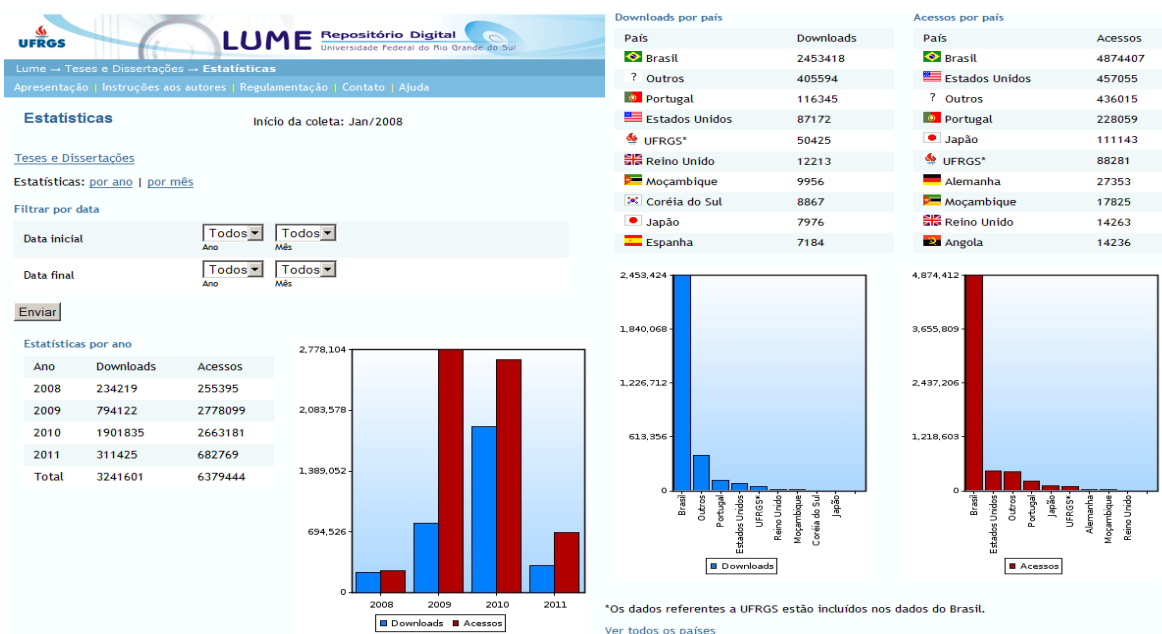


Figura 2. Estadísticas resumidas por comunidad

Como se puede observar en la Figura 2, primero se presentan los datos relativos a los accesos y las descargas de la comunidad, colección o ítem con su distribución anual. Aún es posible visualizar la

distribución mensual de estos datos y limitar el período de las estadísticas de acceso y descarga usando los filtros de año y mes. Al final de la página se presentan los 10 países con mayor número de accesos y descargas, e inmediatamente debajo hay una opción para la visualización de todos los países.

B. ESTADÍSTICAS PARA ÍTEMS DE UN AUTOR O UNA PALABRA CLAVE

Las estadísticas por autor o palabra clave se generan a partir de los ítems recuperados por los índices de autor y palabra clave, disponibles en Lume.

En un primer momento se enumeran los cinco ítems con más descargas y los cinco ítems con menos descargas, como se muestra en la Figura 3. A continuación, se muestran datos muy similares a los datos presentados en las estadísticas de la comunidad.



Figura 3. Cinco trabajos con más y menos descargas.

La información de los cinco trabajos con más y menos descargas son datos importantes para autores y orientadores. De esta manera, ellos pueden constatar cuál de sus trabajos despierta mayor interés público y cuáles un interés menor, y usar esta información como base para decidir qué línea de investigación seguir, los temas que se pueden explorar, entre otros.

C. ESTADÍSTICAS PARA ÍTEMS DE UNA BÚSQUEDA AVANZADA

En la búsqueda avanzada es posible realizar consultas utilizando una combinación de filtros, por texto completo, título, autor, palabra clave y año; también, mostrar las estadísticas producidas a partir de los ítems recuperados por la consulta realizada.

A partir de los resultados recuperados, haciendo clic en el enlace Estadísticas aparecen datos estadísticos muy similares a los presentados para los autores y las palabras clave, pero solamente para los resultados de la búsqueda.

D. ESTADÍSTICAS GENERALES POR COMUNIDAD O COLECCIÓN

Este nuevo módulo permite comparar las estadísticas de acceso entre las comunidades y las colecciones que se encuentran en el mismo nivel jerárquico de Lume. Esta opción se encuentra en el menú de la izquierda, accesible a partir de la autenticación de usuarios autorizados.

Inicialmente, este módulo se dirige a una página que muestra las estadísticas comparadas de todas las comunidades del primer nivel de la jerarquía del Repositorio. Si el usuario está interesado en bajar un nivel en la jerarquía para comparar las estadísticas de subcomunidades/colecciones, es posible hacer clic en el nombre de la comunidad en que desea entrar; por ejemplo, los artículos de revistas. Entonces, a continuación se muestran las estadísticas de todas las subcomunidades/colecciones que están bajo la comunidad elegida.

También es posible filtrar el período por año o mes. Como puede verse en la Figura 4, hay dos representaciones gráficas disponibles con las estadísticas comparadas de las comunidades/colecciones:

- La primera muestra el total de accesos y descargas por comunidad/colección.
- La segunda muestra el promedio de accesos y descarga por ítems de una comunidad/colección. El objetivo de este gráfico es identificar las comunidades/colecciones que, aun con un número menor de ítems depositados, tienen una cantidad significativa de acceso o descarga.

A pesar de todos los datos disponibles, no podemos dejar de considerar que sólo datos cuantitativos pueden ser insuficientes para los análisis comparativos, pues deben tomarse en cuenta, entre otras variables, las diferencias de comunicación científica entre las áreas del conocimiento.

Lume → Estatísticas gerais por comunidade/coleção

[Apresentação](#) | [Instruções aos autores](#) | [Regulamentação](#) | [Contato](#) | [Ajuda](#)

Estatísticas gerais por comunidade/coleção

Início da coleta: Jan/2008

Filtrar por data

Data inicial
Ano Mês

Data final
Ano Mês

[Acervo Fotográfico](#)

[Artigos de Periódicos](#)

[Livros e Capítulos de Livros](#)

[Teses e Dissertações](#)

[Trabalhos de Conclusão de Curso de Especialização](#)

[Trabalhos de Conclusão de Curso de Graduação](#)

[Trabalhos de Eventos](#)

Total

N. de Itens	Downloads:	Acessos
1373	21478	109275
3461	56490	70409
45	9	425
13779	2093349	5048380
959	151630	371243
2636	177462	251368
4057	117482	372100
26310	2617900	6223200

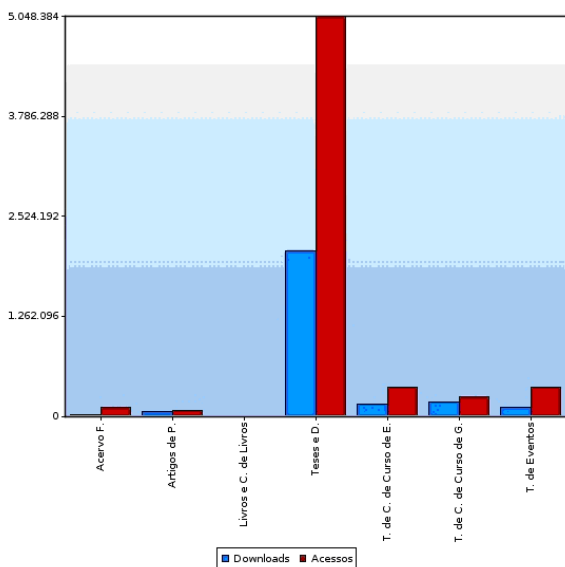


Gráfico do número de downloads e acessos divididos pelo número de itens:

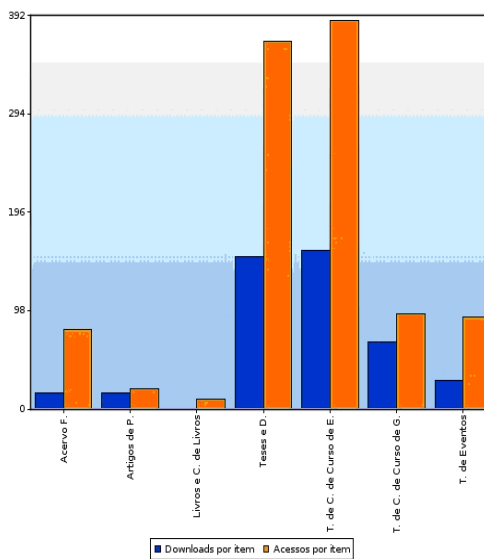


Figura 4. Estadísticas generadas por comunidad/colección.

CONCLUSIONES

Los repositorios institucionales constituyen un importante mecanismo que proporciona amplia visibilidad a la ciencia y devuelve a la sociedad los beneficios con las investigaciones realizadas en el país con recursos públicos, de manera bastante simple y clara; ayudan a aumentar el impacto de la investigación, en

la medida en que se incrementa su visibilidad y accesibilidad, y desempeñan un papel fundamental no sólo en el proceso de comunicación científica, sino también en la preservación y difusión de la producción científica en el marco institucional y nacional.

En este sentido, el trabajo presentó el módulo de estadísticas implantado en Lume - Repositorios Digital de la Universidad Federal de Rio Grande do Sul, que permite la visualización de los datos de accesos y descargas agrupados por comunidad o colección, por un autor o palabra clave, o de un grupo de ítems resultantes de una búsqueda avanzada.

Los beneficios de las estadísticas de uso están relacionados con el hecho de que son eternas. En contraste con las estadísticas de citas basadas en recuentos y *rankings*, las estadísticas de utilización pueden actualizarse en el momento real de uso y son transparentes. Son una medida de la visibilidad de un recurso electrónico, es decir, que se utilizan para que los autores evalúen cuán visible es su trabajo dentro del repositorio [18].

Los datos presentados por las estadísticas permiten a los usuarios y los administradores del sistema analizar la distribución de los accesos y descarga de los ítems del repositorio. A los autores y orientadores les permite, por ejemplo, analizar cuáles son sus obras más consultadas y menos consultadas, así como de qué países son los accesos, el aumento de la visibilidad y el impacto del trabajo que realizan. Para los autores, las estadísticas también pueden considerarse un factor de estímulo para el depósito de sus trabajos en el repositorio institucional, por la información con valor añadido que ofrecen. Para los gestores y las agencias de fomento son una fuente considerable de datos para la planificación, la toma de decisiones, el acompañamiento y la evaluación de la investigación en la institución y en el país.

Además de las informaciones proporcionadas por este módulo, hay otras tantas en los registros de *logs* de acceso del sistema que permite estudios más refinados sobre la visibilidad, el impacto, y el comportamiento de los usuarios en el uso de Lume, entre otros. Todavía queda mucho por explorar y estudiar sobre los análisis que se pueden hacer con los datos recogidos y cómo pueden contribuir para la gestión de la investigación, de las instituciones y de sus repositorios.

REFERENCIAS

- [1] DSpace. [Online]. Available: <http://www.dspace.org/>
- [2] B. S. Campelo, "Pesquisas em andamento," en: Fontes de informação para pesquisadores e profissionais. Belo Horizonte: UFMG, 2000, Pp. 49-54.
- [3] A. M. N. Marques and S. R. da S. Maio, "Repositórios institucionais." 2007. Available: <http://repositoriosdigitais.web.simplesnet.pt/PDF'S/Artigo%20%20Repositorios%20Institucionais.pdf>

-
- [4] F. Sarmiento, A. Miranda, A. Baptista, I. Ramos, "Algumas considerações sobre as principais declarações que suportam o movimento Acesso Livre." 2005. Trabalho apresentado no 9º World Congress on Health Information and Libraries, Salvador. Available: <http://repositorium.sdum.uminho.pt/handle/1822/4282>.
- [5] A. J. Meadows, *A comunicação científica*. Brasília: Briquet de Lemos, 1999.
- [6] P. Bourdieu, "O campo científico," en R. Ortiz, *Pierre Bourdieu: Sociologia*. São Paulo: Ática, 1994. pp. 122-155.
- [7] S. M. S. Costa, "Filosofia aberta, modelos de negócios e agências de fomento: elementos essenciais a uma discussão sobre o acesso aberto à informação científica," *Ciência da Informação, Brasília*, vol. 35, no. 2, pp. 39-50, maio/ago. 2006.
- [8] R. Saraiva, "Plano de comunicação para repositórios institucionais," [Portugal]: [s.n.], 2008. [Online]. Available: http://projecto.rcaap.pt/index.php?option=com_remository&Itemid=2&func=startdown&id=97&lang=pt
- [9] S. P. M. Mueller, "Políticas de fomento à consolidação da capacidade de pesquisa no Brasil: Estudo para avaliação do Programa Pós-Doutoral no Exterior mantido pela CAPES," Universidade de Brasília, 2004, Relatório de pesquisa em andamento.
- [10] S. P. M. Mueller, "A comunicação científica e o movimento de acesso livre ao conhecimento," *Ciência da Informação, Brasília*, vol. 35, no. 2, pp. 27-38, maio/ago. 2006. [Online]. Available: <http://revista.ibict.br/ciinf/index.php/ciinf/article/view/826>
- [11] S. Goulart and C. A. Carvalho, "O caráter da internacionalização da produção científica e sua acessibilidade restrita," *Revista de Administração Contemporânea, Curitiba*, vol. 12, no. 3, pp. 835-853, jul./set. 2008.
- [12] G. Whitney, "Patterns of authorship in major bibliographic databases: the European region," *Scientometrics*, vol. 26, no. 2, pp. 275-292, 1993.
- [13] J. Houghton and P. Sheehan, "Estimating the Potential Impacts of Open Access to Research Findings," *Economic Analysis and Policy*, vol. 39, no. 1, pp. 127-142, Mar. 2009. [Online]. Available: <http://eap-journal.com/download.php?file=696>
- [14] Corporation for National Research Initiatives, "The Handle System," 2008. [Online]. Available: <http://www.handle.net/>

- [15] Dublin Core, “Dublin Core Metadata Initiative,” 2010. [Online]. Available: <http://www.dublincore.org/documents/usageguide>
- [16] The Open Archives Initiative Protocol for Metadata Harvesting, 2004. [Online]. Available: <http://www.openarchives.org/OAI/openarchivesprotocol.html>
- [17] C. Merk and N. K. Windisch, “JISC Usage Statistics Review: final report,” 2008. [Online]. Available: http://ie-repository.jisc.ac.uk/250/1/Usage_Statistics_Review_Final_report.pdf

SOBRE LOS AUTORES

C. G. Pavão

Possui graduação em Biblioteconomia e Documentação (1986) mestrado em Comunicação e Informação (2010) e é Doutoranda em Comunicação e Informação pela Universidade Federal do Rio Grande do Sul (UFRGS). É bibliotecária da UFRGS e desenvolve suas atividades junto à Divisão de Sistemas de Informação do Centro de Processamento de Dados da UFRGS. Tem experiência em bases de dados, automação de bibliotecas, comunicação científica e repositórios digitais.

S. E. Caregnato

Possui graduação em Biblioteconomia pela Universidade Federal do Rio Grande do Sul (UFRGS, 1987), mestrado em Information Management pela University of Sheffield (1992) e doutorado em Information Studies pela University of Sheffield (2000). É Bolsista de Produtividade em Pesquisa do CNPq - Nível 2 e professora adjunta da Faculdade de Biblioteconomia e Comunicação (FABICO) e do Programa de Pós-Graduação em Comunicação e Informação (PPGCOM) da UFRGS. Tem experiência na área de Ciência da Informação, atuando principalmente nos seguintes temas: comunicação científica, periódicos eletrônicos, ciëntometria, objetos de aprendizagem, telecentros comunitários e inclusão digital.

Z. Horowitz

Possui graduação em Matemática, especialização em Engenharia e Gerência de Software pela Universidade Federal do Rio Grande do Sul (UFRGS). É diretora da Divisão de Sistemas de Pesquisa do Departamento de Sistemas de Informação do Centro de Processamento de Dados da UFRGS desde 2004. Possui interesse nas seguintes áreas: repositórios digitais, engenharia e gerência de software e ciência e gestão da informação.

J. S. B. da Costa

Possui graduação em Biblioteconomia e Documentação (1982) e especialização em Gestão de Bibliotecas Universitárias (2009) pela Universidade Federal do Rio Grande do Sul (UFRGS). É bibliotecária da UFRGS e desenvolve suas atividades junto à Divisão de Sistemas de Informação do Centro de Processamento de Dados da UFRGS. Tem experiência em bases de dados, automação de bibliotecas e repositórios digitais.

M. K. Ferreira

Possui graduação em Ciência da Computação (2008) e é mestranda do Programa de Pós-Graduação em Computação no Instituto de Informática da Universidade Federal do Rio Grande do Sul (UFRGS). É analista de sistemas no Centro de Processamento de Dados da UFRGS desde 2009. Possui interesse nas seguintes áreas: programação paralela e distribuída, arquiteturas multicore, repositórios digitais, engenharia de software e complexidade de problemas (estudo de melhores algoritmos para a resolução de problemas).