

**Estadísticas de uso de repositorios institucionales: ¿un indicador eficaz de comparación?:** ideas para la ponencia en las Jornadas OS-REPOSITORIOS del próximo diciembre en Madrid.

## GRÁFICO COMPARATIVO DE VISITAS/DESCARGAS EN REPOSITORIOS ESPAÑOLES.

Por el momento sólo **unos pocos** repositorios españoles ofrecen estadísticas de uso: UPCommons, RUA, e-archivo uc3m, e-Espacio UNED, Digital.CSIC. Modelos distintos (diferentes softwares, ePrints, DSpace, Fedora, diferencias entre plataformas con mismo software de base). [PEDIR ASESORAMIENTO A LOS COMPAÑEROS]

Consideraciones generales sobre el **análisis previo** de las estadísticas para garantizar su homogeneidad (ver notas más abajo). Tratamiento previo de los datos (modelo). Desigual cumplimiento, ausencia de criterios generales. Proyecciones anuales (irreales), extrapolaciones dudosas [PEDIR etc].

Ranking Web de Repositorios: varios elementos de comparación en el modelo, aunque no estadísticas de acceso por el momento. Propuesta del Laboratorio de Cibermetría: analizador estadístico común para garantizar una salida equiparable (ver notas más abajo).

Indicadores normalizados para una comparativa eficaz: visitas y descargas por registro, visitas y descargas por usuario del sistema.

En el momento presente, se trata de una herramienta básicamente orientada a la monitorización de la evolución a nivel interno, poco apta para la evaluación comparativa. Posiblemente iniciativas como DRIVER puedan reorientar este estado de cosas. Coordinación nacional-plataforma Recolecta.

[COMENTAR CON ISIDRO, VICTORIA Y SI ES POSIBLE, CON ALGUIEN DE ePrints (¿UCM?)]

-----

(fuente: RECERCAT, <http://www.recercat.net>)

### Notas

1. Una consulta a un documento se define como un acceso a su página de metadatos, o como la descarga directa de este documento cuando ésta no es a través de su propia página de metadatos.
2. Los accesos automáticos de robots, tanto de aquellos que consultan el fichero *robots.txt* como otros detectados por nosotros, no se contabilizan como consultas. También se excluyen los accesos fraudulentos (muchos accesos desde una misma IP) y los accesos realizados desde direcciones del CIESA y del CBUC.
3. Para determinar las consultas por país se usan las librerías GeoIP.

(fuente: Ranking of World Repositories July'08, <http://repositories.webometrics.info/>)

### **Proposal to webmasters for 2009 edition**

Usage data could be very useful for the Ranking Web of World Repositories as it has been suggested by most of the academic community. It is difficult to obtain data from every repository and in fact most of the figures are not comparable because there is no standard yet about statistics collection. We suggest using a common, simple and free system to solve, at least partially, this problem. We ask to webmasters to install Google Analytics (<http://www.google.com/analytics>) at their websites. We will then be able to kindly ask you for the statistics collected during the period from July 1st -December 31st so this data can be used for the January 2009 edition of the Ranking.

(fuente: Chris Keene, Sussex University, UK, <http://www.nostuff.org/ircount/info.php>)

### **Lies, more lies and statistics**

It will be silly to base success on - or use as a measure of performance - the number of records in an IR. This site has the potential to encourage people to compare IRs by record count and conclude one IR is doing 'better' than another. The number of records really isn't a good measure of anything. IRs are primarily about full text documents and articles, not simply records. Some organisations will have historic publication databases which they may put in to their IR. Plus IRs have different criteria, those with a wide remit such as including learning objects, images, even newsletters, may well have higher counts than others though it says nothing as to which has the largest pool of research.

What's a more successful IR: one with many records, or one with a few but each with a publicly available full text file of the research? One with many basic records of poor quality, or one with only a few that are complete and accurate? One with many records for the Library's Annual Report, Newsletter, powerpoint presentations etc, or one with less items but each being a piece of high quality research that has been published in a highly regarded journal? This website can do nothing more than what it does. That is show the number of records in each IR. Use caution when drawing conclusions from this data.

<http://www.nostuff.org/ircount/table.php?country=es>