

Contenidos de I+D en Internet: Mitos y Leyendas

Isidro F. Aguillo
CINDOC-CSIC
isidro@cindoc.csic.es

La creciente popularización de Internet ha dado lugar a una cierta banalización de la misma en los medios de comunicación. Una consecuencia negativa de este proceso es la aparición de una amplia serie de datos incorrectos o falsos sobre los contenidos de la red que podrían restar valor en su utilización por científicos y tecnólogos. Con el apoyo de algunos datos empíricos se discuten algunos de los tópicos inexactos y se analiza el impacto potencial y la penetración real de la Web en el entorno académico y de investigación.

Internet es una red fundamentalmente comercial

Los procesos de globalización y la emergencia de la llamada nueva economía han llevado a suponer que el carácter fundamental de la red es básicamente mercantilista y que el comercio electrónico es el motor fundamental del crecimiento de Internet.

Dicha apreciación no es un reflejo real de la situación ya que, como demuestran los datos (Tabla I), un porcentaje no desdeñable del WWW tiene un carácter netamente académico (20-25%). Más difícil resulta precisar el peso de la I+D, pero considerando tanto la contribución pública como la privada se puede estimar que alrededor del 10% de los contenidos tienen dicho carácter en la actualidad.

PAIS	Dominio	Páginas	Subdominio	Páginas	%
JAPON	jp	16,276.00 9	ac.jp	3,065.568	18.83
REINO UNIDO	uk	8,696.572	ac.uk	2,264.384	26.04
AUSTRALIA	au	4,766.177	edu.au	1,411.770	29.62
COREA DEL SUR	kr	3,915.505	ac.kr	1,188.885	30.36
TAIWAN	tw	3,131.568	edu.tw	1,345.589	42.97
POLONIA	pl	1,967.533	edu.pl	163.527	8.31
AUSTRIA	at	1,840.582	ac.at	540.371	29.36
BELGICA	be	1,151.946	ac.be	264.385	22.95
CHINA	cn	3,039.371	edu.cn	338.111	11.12
SUDAFRICA	za	861.079	ac.za	106.016	12.31
NUEVA ZELANDA	nz	731.494	ac.nz	132.134	18.06
ISRAEL	il	557.137	ac.il	167.098	29.99
ARGENTINA	ar	738.083	edu.ar	64.334	8.72

Tabla I. Porcentaje de páginas académicas respecto del total según el buscador FAST (16 Feb. 2000)

Internet ofrece contenidos de poca calidad

Internet ofrece un gran volumen de contenidos, que previsiblemente ya han superado los 1.500 millones de páginas a principios del año 2000 y que podrían alcanzar la cifra de 10.000 millones antes del 2005. Posiblemente ni siquiera esta cifra marque la asíntota en el crecimiento explosivo del Web.

Sin embargo, se ha generalizado la crítica sobre la falta de calidad en dichos contenidos, aseveración muy relacionado con el supuesto gran volumen de sedes “conflictivas”. Sin embargo, los datos reales (Steve Lawrence; Lee Giles, 1999. Accessibility and Distribution of Information on the Web. Nature, 400: 107-109) muestran que sólo el 1,5% tiene ese carácter.

Muy al contrario, en los últimos años se está incrementando notablemente el tamaño de la llamada “Internet invisible”, el conjunto de información que aún siendo accesible desde la red a través de pasarelas no puede ser indizado estructuralmente por los motores de búsqueda. Son recursos que constituyen auténticas “islas de calidad” y cuya importancia es cada vez mayor. Podemos incluir aquí los miles de bases de datos y catálogos de bibliotecas ya disponibles.

Incluir Figura MEDLINE

También se está produciendo un cambio de paradigma en los procesos de comunicación científica, lo que ha incrementado notablemente la presencia de revistas científicas en la red, incluyendo versiones electrónicas de algunas de las de mayor impacto. Dichas publicaciones forman también parte de esa Internet invisible.

Incluir Figura HIGHWIRE

Además muchas comunidades de investigadores comparten sus trabajos en la red antes y después de su publicación formal para incrementar y agilizar su distribución y facilitar la evaluación de los mismos por parte de sus colegas.

Incluir Figura E-PRINT

Posición	Dominio	Institución	“hosts”	%
3	uclm.es	UNIVERSIDAD DE CASTILLA LA MANCHA	27722	5.68%
6	uned.es	UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA	13260	2.72%
7	upv.es	UNIVERSITAT POLITECNICA DE VALENCIA	12501	2.56%
9	ehu.es	UNIVERSIDAD DEL PAÍS VASCO	10569	2.17%
10	upc.es	UNIVERSITAT POLITÈCNICA DE CATALUNYA	9329	1.91%
11	upm.es	UNIVERSIDAD POLITÈCNICA DE MADRID	8901	1.82%
12	uv.es	UNIVERSITAT DE VALENCIA	8717	1.79%
13	uclm.es	UNIVERSIDAD COMPLUTENSE DE	8087	1.66%

		MADRID		
14	csic.es	CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS	7192	1.47%
15	uniovi.es	UNIVERSIDAD DE OVIEDO	6668	1.37%
16	ub.es	UNIVERSITAT DE BARCELONA	6529	1.34%
17	ugr.es	UNIVERSIDAD DE GRANADA	6199	1.27%
18	uab.es	UNIVERSITAT AUTÓNOMA DE BARCELONA	5970	1.22%
19	um.es	UNIVERSIDAD DE MURCIA	5710	1.17%
21	uam.es	UNIVERSIDAD AUTÓNOMA DE MADRID	5199	1.07%
22	unizar.es	UNIVERSIDAD DE ZARAGOZA	5176	1.06%
23	uma.es	UNIVERSIDAD DE MÁLAGA	5071	1.04%
25	uva.es	UNIVERSIDAD DE VALLADOLID	4430	0.91%
26	usc.es	UNIVERSIDADE DE SANTIAGO DE COMPOSTELA	4379	0.90%
27	us.es	UNIVERSIDAD DE SEVILLA	4155	0.85%
28	unileon.es	UNIVERSIDAD DE LEÓN	4009	0.82%
30	usal.es	UNIVERSIDAD DE SALAMANCA	3641	0.75%
31	uc3m.es	UNIVERSIDAD CARLOS III DE MADRID	3570	0.73%
33	upf.es	UNIVERSITAT POMPEU FABRA	3339	0.68%
35	ulpgc.es	UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA	3280	0.67%
~200		TOTAL I+D	241161	49.43%
16868	es	TOTAL ESPAÑA	487851	

Penetración de Internet en las principales instituciones de I+D españolas respecto del total (RedIRIS, Enero 2000).

Internet es caótico y poco organizado

Los contenidos disponibles en Internet son publicados en la red por los propios productores de los mismos bajo la responsabilidad de las instituciones hospedadoras y buscando el cumplimiento de sus fines y objetivos según sus características propias, necesidades y posibilidades.

De forma coyuntural la intermediación que está organizando los recursos depositados en la red tiene un fuerte carácter comercial y, por tanto, importantes sesgos.

Sin embargo, la creciente necesidad de especialización está devolviendo el protagonismo a los actores institucionales, de forma que los actuales directorios se organizan según las clasificaciones normalizadas y de acuerdo a criterios documentales.

Insertar fichero BUBL

Asimismo, los sistemas de gestión de las bases de datos soportan operadores avanzados de tipo “booleano”, con amplias posibilidades para delimitar las búsquedas por campos. La irrupción de sistemas super-avanzados terminará por cambiar la percepción de falta de organización de los recursos en la red.

1	Albert Einstein	117.512
2	Isaac Newton	44.768
3	James Clerk Maxwell	6.375
4	Niels Bohr	13.039
5	Werner Heisenberg	7.478
6	Galileo Galilei	19.676
7	Richard Feynmann	289
8	Paul Dirac	2.609
9	Erwin Schrodinger	1.076
10	Ernest Rutherford	3.713

Los diez físicos más importantes según la revista inglesa **Physics World** (1999) y el número de veces que son citados en la Web (FAST, Feb. 2000)

Internet es lento y complejo de manejar

Los actuales problemas técnicos de la red no son de carácter estructural, de forma que los avances para su resolución en los últimos años han sido considerables.

La saturación del tráfico se eliminará con la instalación de las nuevas líneas de gran ancho de banda y la demanda, que ha estado creciendo exponencialmente, alcanzará un comportamiento asintótico a medio plazo.

Desde un punto de vista práctico, Internet tiene mucho más que ver con información que con informática, de forma que la actual generación de clientes universales es más transparente para el usuario final y el auge de los sistemas hipermedios basados en el World Wide Web así lo demuestra.

Insertar Figura CYBERMETRICS

Asimismo se ha incrementado notablemente el nivel de formación de los usuarios finales y la cantidad y calidad de la documentación de apoyo.

Internet distribuye la información de forma gratuita

La información publicada en la red tiene un costo considerable, que se financia tanto desde el sector público como desde el privado, fundamentalmente a través de publicidad. El usuario final paga, de una forma u otra, la generación y publicación de los recursos. El coste de la distribución electrónica es insignificante comparado con lo que representa la producción de información, que en muchos casos se realiza independientemente de su destino final.

Figura PORTAL_AIRCENTER

La comunidad de científicos y tecnólogos del sector público tienen el reto de alcanzar de forma más amplia y general a la sociedad a través de cuyos ciudadanos se financia su actividad.

Posición	Institución	dominio	páginas	%
1	UNIVERSIDAD POLITÉCNICA DE MADRID	upm.es	52.399	2,31%
2	UNIVERSITAT POLITÉCNICA DE CATALUNYA	upc.es	49.881	2,20%
3	UNIVERSITAT DE BARCELONA	ub.es	48.884	2,16%
4	UNIVERSITAT JAUME I	uji.es	41.993	1,85%
5	UNIVERSITAT POLITECNICA DE VALENCIA	upv.es	39.519	1,74%
6	UNIVERSITAT AUTÓNOMA DE BARCELONA	uab.es	39.222	1,73%
7	CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS	csic.es	25.839	1,14%
8	UNIVERSIDAD DE CANTABRIA	unican.es	23.411	1,03%
9	UNIVERSIDAD DE ZARAGOZA	unizar.es	22.743	1,00%
10	UNIVERSIDAD DEL PAÍS VASCO	ehu.es	18.881	0,83%
11	UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA	ulpgc.es	18.127	0,80%
12	UNIVERSITAT DE VALENCIA	uv.es	17.412	0,77%
13	UNIVERSIDAD CARLOS III DE MADRID	uc3m.es	15.913	0,70%
14	UNIVERSIDAD DE OVIEDO	uniovi.es	15.872	0,70%
15	UNIVERSIDADE DE SANTIAGO DE COMPOSTELA	usc.es	14.617	0,65%
16	UNIVERSIDAD DE GRANADA	ugr.es	14.530	0,64%
17	UNIVERSIDAD AUTÓNOMA DE MADRID	uam.es	13.632	0,60%
18	UNIVERSITAT ROVIRA I VIRGILI	urv.es	13.625	0,60%
19	UNIVERSIDADE DA CORUÑA	udc.es	13.005	0,57%
20	UNIVERSIDAD DE VALLADOLID	uva.es	12.927	0,57%
21	UNIVERSIDAD DE SEVILLA	us.es	12.902	0,57%
22	UNIVERSITAT POMPEU FABRA	upf.es	12.814	0,57%
23	UNIVERSIDAD COMPLUTENSE DE MADRID	ucm.es	12.682	0,56%
24	UNIVERSIDADE DE VIGO	uvigo.es	11.237	0,49%
25	UNIVERSIDAD DE SALAMANCA	usal.es	10.988	0,48%
	TOTAL ESPAÑA	es	2.266.141	

Principales instituciones de I+D en España según tamaño (número de páginas. FAST 15 Febrero de 2000). Las 25 primeras suponen el 25,29% del total.

La posibilidad de que se establezcan monopolios de hecho en el área de contenidos es preocupante tanto por su propio carácter, como por el peligro real de colonización cultural que el predominio del inglés como “lingua franca” puede generar.

Insertar Figura INKTOMI

Solo un intenso programa de transferencia de recursos a formato electrónico podría paliar dicha situación, por lo que se deben generar contenidos propios que reflejen nuestra idiosincrasia, riqueza y diversidad cultural y la actividad de nuestro sistema ciencia – tecnología - industria.