



Buscadores académicos

José Luis Ortega
Laboratorio de Cibermetría
jortega@orgc.csic.es



La Web: cambio de paradigma

- Recuperación de información pre-Web
 - Bases de datos y directorios
 - Usuarios especializados
 - Contenidos especializados
 - Documentos estructurados, homogéneos, de alta calidad



La Web: cambio de paradigma

- Recuperación de información en la Web
 - Buscadores
 - Usuarios generalistas y bajo nivel
 - Contenidos generales
 - Documentos desestructurados, formatos diversos, heterogéneos



La Web: cambio de paradigma

- Buscadores
 - Simplificación de interfaces
 - Enorme cantidad de resultados (ranking)
 - Generalistas



La Web: cambio de paradigma

- Buscadores académicos
 - Libre acceso
 - Especializados en documentos científicos (artículos, libros, comunicaciones a congresos, patentes, etc.)
 - Instrumentos de evaluación (citas)
 - Valor añadido (indicadores, agregados, comparativas, etc.)

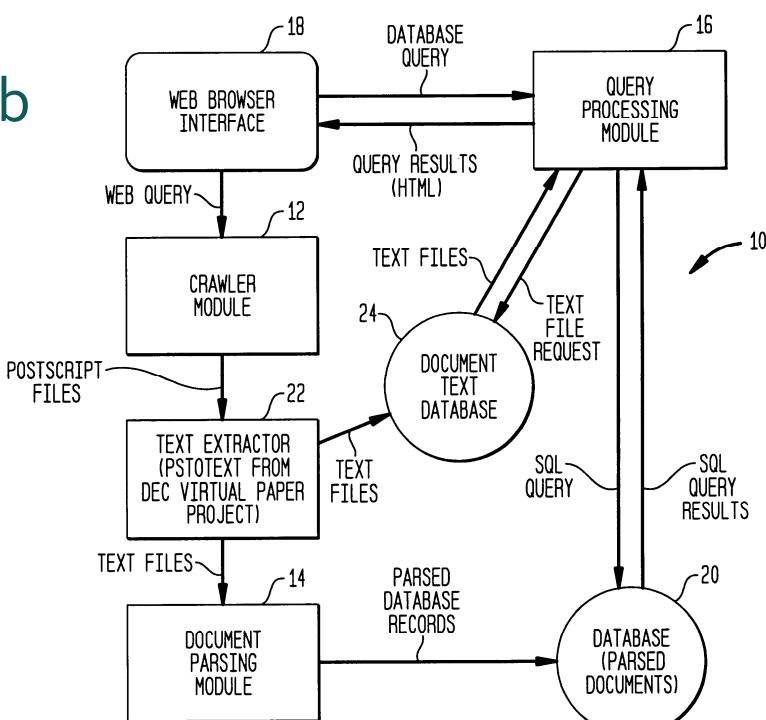


CiteSeer^x_β

- Primer buscador académico
 - CiteSeer: 1997 en NEC (Lawrence, Lee Giles y Bollacker)
 - CiteSeer^x: 2008 en UPenn (Lee Giles)
- Especializado en Informática y Ciencias de la Información
- Primer índice autónomo de citas
- Biblioteca digital

CiteSeer^x_β

- Índice autónomo de citas:
 - Rastreo (*crawling*) de la Web académica (.pdf, .ps y .zip)
 - Almacenamiento de documentos (txt) en su biblioteca digital
 - Extracción (*parsing*):
 - metadatos sobre el artículo
 - Referencias bibliográficas e información contextual de la cita



(US6289342 B1, 2001)



CiteSeer^x_β

- Cobertura
 - 3,3 millones de documentos (BASE, 2014)
 - 300 mil perfiles de autores
 - Especializado en informática y ciencias de la información



CiteSeer^x_β

- Búsqueda
 - general de documentos
 - de autores
 - de títulos
 - avanzada
- Sólo 500 registros
- Ranking
 - Citas
 - Relevancia
 - Fecha
 - Recencia

Functions from a set to a set (1989)

by Czesław Byliński

Venue: Journal of Formalized Mathematics

Citations: 1007 - 23 self

[Save to List](#)[Add to Collection](#)[Correct Errors](#)[Monitor Changes](#)[Cached](#)[Download Links](#)[\[mizar.org\]](#)[\[markun.cs.shinshu-u.ac.jp\]](#)[\[www.cs.ualberta.ca\]](#)[\[mizar.uwb.edu.pl\]](#)[Summary](#)[Active Bibliography](#)[Co-citation](#)[Clustered Documents](#)[Version History](#)

Abstract

function from a set X into a set Y , denoted by "Function of X,Y ", the set of all functions from a set X into a set Y , denoted by $\text{Funcs}(X,Y)$, and the permutation of a set (mode Permutation of X , where X is a set). Theorems and schemes included in the article are reformulations of the theorems of [1] in the new terminology. Also some basic facts about functions of two variables are proved.

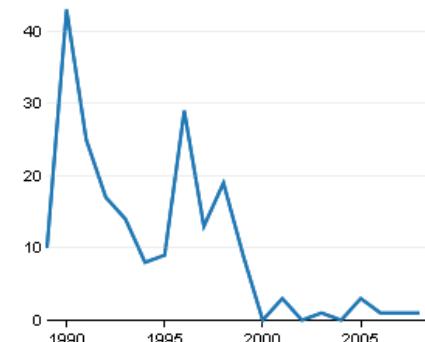
Citations

- 1276 Tarski Grothendieck set theory - Trybulec - 1989
- 1252 Functions and their basic properties - Byliński - 1990
- 1178 Properties of subsets - Trybulec - 1989
- 982 Relations and their basic properties - Woronowicz - 1990
- 721 Some basic properties of sets - Byliński - 1989
- 449 Relations defined on sets - Woronowicz - 1990
- 426 Partial functions - Byliński - 1990

BibTeX

```
ARTICLE(Byliński89functionsfrom,  
        author = {Czesław Byliński},  
        title = {Functions from a set to a set},  
        journal = {Journal of Formalized  
Mathematics},  
        year = {1989},  
        volume = {1},  
        pages = {1--2}  
)
```

Years of Citing Articles





- Errores en la extracción de información (*parsing*)

Abstract

function from a set X into a set Y , denoted by “Function of X,Y ”, the set of all functions from a set X into a set Y , denoted by $\text{Funcs}(X,Y)$, and the permutation of a set (mode Permutation of X , where X is a set). Theorems and schemes included in the article are reformulations of the theorems of [1] in the new terminology. Also some basic facts about functions of two variables are proved.

Summary. The article is a continuation of [1]. We define the following concepts: a function from a set X into a set Y , denoted by “Function of X,Y ”, the set of all functions from a set X into a set Y , denoted by $\text{Funcs}(X,Y)$, and the permutation of a set (mode Permutation of X , where X is a set). Theorems and schemes included in the article are reformulations of the theorems of [1] in the new terminology. Also some basic facts about functions of two variables are proved.

Abstract

B-trees have become, de facto, a standard for file organization. File indexes of users, dedicated database systems, and general-purpose access methods have all been proposed and implemented using B-trees. This paper reviews B-trees and shows why they have been so successful. It discusses the major variations of the B-tree, especially the B+-tree,

B-trees have become, de facto, a standard for file organization. File indexes of users, dedicated database systems, and general-purpose access methods have all been proposed and implemented using B-trees. This paper reviews B-trees and shows why they have been so successful. It discusses the major variations of the B-tree, especially the B⁺-tree, contrasting the relative merits and costs of each implementation. It illustrates a general purpose access method which uses a B-tree.

- Fallos en la captura de citas

Results 1 - 10 of 1.328.451

Next 10 →

[CITATION] Hassenblatt B., Introduction to the Modern Theory of Dynamical Systems "

by A Katok , 1996

"...."

- Cited by 666 (20 self) - Add to MetaCart

Anatole Katok and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza.

[CITATION] The B-Book: Assigning programs to meanings "

by J R Abrial , 1996

"...."

- Cited by 644 (6 self) - Add to MetaCart

J.-R. Abrial. *The B-Book: Assigning Programs to Meanings*. Cambridge University Press, 1996.

[CITATION] Recursively generated B-spline surfaces on arbitrary topological meshes "

by E Catmull, J Clark - *Computer-Aided Design* , 1978

"...."

- Cited by 541 (0 self) - Add to MetaCart

E. Catmull and J. Clark. Recursively generated B-spline surfaces on arbitrary topological meshes. *Computer Aided Design*, 10:350–355, 1978.

[CITATION] Yuan,B.(2000).Fuzzy Sets and Fuzzy Logic: Theory and Applications "

by G J Klir

"...."

- Cited by 677 (2 self) - Add to MetaCart

G.J. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall, 1995.



CiteSeer^x_β

- Mala asignación de fechas (3,5%)

[Documents](#) [Authors](#) [Tables](#) [MetaCart](#) [Sign up](#) [Log in](#)

CiteSeer^x_β

year:[0 TO 1900]

Include Citations

Search

[Advanced Search](#)

Results 1 - 10 of 72.024

[Next 10 →](#)

Tools

Sorted by:

Relevance

Try your query at:

[Scholar](#) [Yahoo!](#) [Ask](#)

[Bing](#) [CSB](#) [Libra](#)

The Preprocessor

by unknown authors , 1580

"... ♦Initial step prior to analysis by the compiler. ♦The text of a C++ program is scanned and directives that begin with a hash mark are processed. ♦Major categories are: #include <filename> #define name value #undef name #if test #ifdef name #ifndef name ..."

[Abstract](#) - [Add to MetaCart](#)

human genetic linkage analysis

by E. A. Thompson , 709

"... Abstract: This paper summarizes and discusses the methodological research in human genetic linkage analysis, leading up to and following from ..."

[Abstract](#) - [Add to MetaCart](#)

CONTENTS

by unknown authors , 1900

"... 1900-1993 E=A*EF*[1- C/100] ..."

[Abstract](#) - [Add to MetaCart](#)

Wolfgang Amadeus Mozart

by David H. D. Warren , 1756

"... I dedicate this modest work to a most outstanding ..."

[Abstract](#) - [Add to MetaCart](#)

American Bruhn / Higher Behavioral Education Scientist as Empowerment Higher Education as Empowerment The Case of Palestinian Universities

by Christa Bruhn , Christa Bruhn , 1125

"... Palestinian universities have been a dynamic force across Palestinian communities since their inception under Israeli occupation in the early 1970s. What began as initiatives of private families or religious foundations to provide local access to higher education soon grew into a widespread presence ..."

[Abstract](#) - [Add to MetaCart](#)



CiteSeer^x_β

- Inconsistencia en el cálculo de citas

[Mining Quantitative Association Rules in Large Relational Tables](#)

by Ramakrishnan Srikant, Rakesh Agrawal , 1996

"... We introduce the problem of mining association rules in large relational tables containing both quantitative and categorical attributes. An example of such an association might be "10% of married people between age 50 and 60 have at least 2 cars". We deal with quantitative attributes by finepartitio ..."

Abstract - [Cited by 299 \(2 self\)](#) - [Add to MetaCart](#)

Mining Quantitative Association Rules in Large Relational Tables (1996)

by Ramakrishnan Srikant , Rakesh Agrawal

[Citations: 304 - 2 self](#)

Mining Quantitative Association Rules in Large Relational Tables (1996)

by Ramakrishnan Srikant, Rakesh Agrawal

[Add To MetaCart](#)



CiteSeer^x_β

- 150.000 documentos extraídos (2014)
 - 14,2 % correspondía a documentos eliminados
 - 1,6% no tenían título
 - 45% no tenía fecha de publicación
 - 32% no tenía fuente de publicación (revista, congreso, etc.)
 - 36,7% asignados a un autor



CiteSeer^x_β

- Conclusiones:
 - Pequeño tamaño y especializado
 - Graves problemas en la extracción de datos
 - Afecta a la recuperación
 - Influye como herramienta de evaluación
 - Biblioteca digital



- Nace en 2004 de la mano de Anurag Acharya
- Misión: el principal punto de acceso a la literatura científica de forma gratuita y universal
- Perfección del índice autónomo de citas
- Amplia los documentos académicos (libros, tesis, informes, etc.)



- Emerge de la arquitectura de Google
- Rastrea webs de universidades, editoriales científicas y repositorios
- Condición:
 - artículos científicos con resumen
 - metadatos legibles por su robot
- Excluye:
 - noticias, artículos en magacines, recensión de libros, etc.
 - materiales de más de 5MB, que es enviado a Google Books





- Gestión de duplicados:
 - toma datos una sola vez de cada documento
 - siempre de fuentes autorizadas (i.e. plataformas editoriales)
 - El resto se incluye como versiones del documento original o publicado

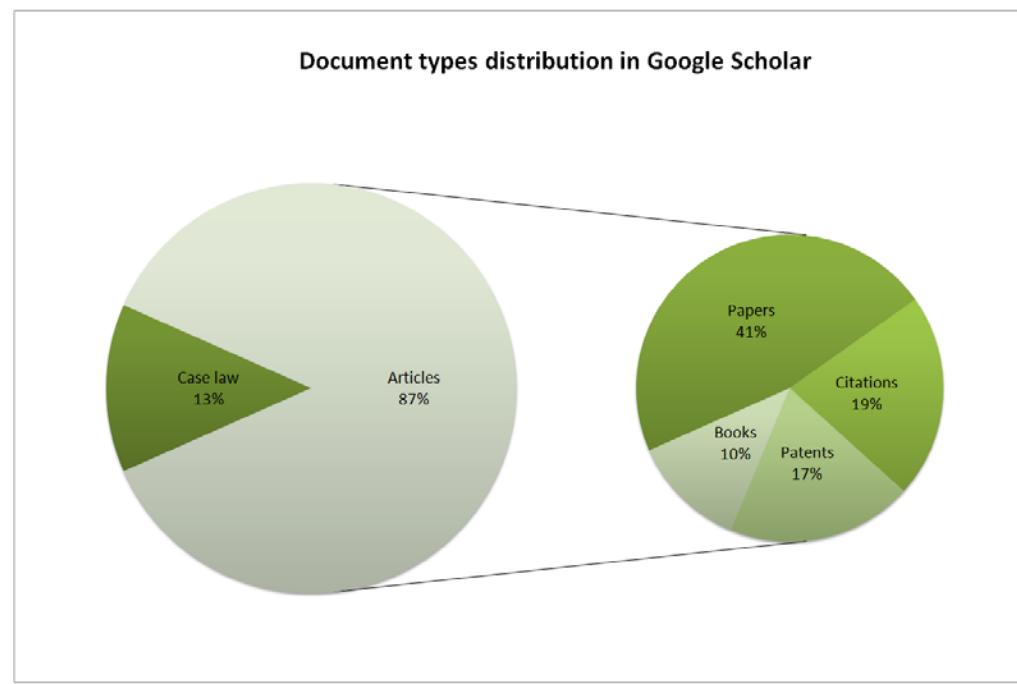


Google scholar

- Cobertura

- 87 M. (Aguillo, 2012)
- 100 M. (Khabsa & Giles, 2014)
- 109 M. (Ortega, 2014)
- 160 M. (Orduña et al., 2014)

Document type	total	%
Academic papers	44,403,310	46.83
Citations	20,394,540	21.51
Patents	18,553,865	19.57
Books	11,467,605	12.09
Total	94,819,320	100





- Distribución por tipo de fuentes

Sources type	Total	%	articles	%	citations	%
Publishers	30,822,200	41.60	28,711,300	58.84	2,110,900	10.42
Google Services	16,630,000	22.45	4,000,000	8.20	7,600,000	37.51
Thematic repositories	12,550,000	16.94	6,595,000	13.52	5,955,000	29.39
Institutional repositories	8,739,000	11.80	7,108,400	14.57	1,630,600	8.05
Bibliographic services	5,263,100	7.10	2,300,600	4.71	2,962,500	14.62
Online journals	80,500	.11	80,500	.16	0	.00
Total	74,084,800	100.00	48,795,800	100.00	20,259,000	100.00

- Distribución por fuentes

Publisher	Source Type	Total	%	articles	%	citations	%
Google Books	Google service	11,600,000	12.3	4,000,000	5.4	7,600,000	37.3
ncbi.nlm.nih.gov (PubMed Central)	Thematic repository	7,310,000	7.7	2,830,000	3.8	4,480,000	21.9
Google Patents	Google service	5,030,000	5.3	0	.0	0	.0
Elsevier	Publisher	4,380,000	4.6	4,420,000	5.9	-40,000	-.2
Springer	Publisher	3,800,000	4.0	3,500,000	4.7	300,000	1.5
.edu (TLD university)	Institutional repository	3,340,000	3.5	2,560,000	3.4	780,000	3.8
Science	Publisher	3,180,000	3.3	2,960,000	3.9	220,000	1.1
inist.fr (INIST databases)	Bibliographic database	3,010,000	3.2	1,450,000	1.9	1,560,000	7.6
ieeexplore.ieee.org (IEEE Xplore)	Publisher	2,980,000	3.1	2,760,000	3.7	220,000	1.1
Wiley Online Library	Publisher	2,810,000	2.9	2,710,000	3.6	100,000	.5
All publishers		94,819,320		74,424,780		20,394,540	



- **Case law**
 - Creado en 2009 como *Court opinions*
 - principales resoluciones de los tribunales de justicia más importantes de Estados Unidos
 - 29 M. de documentos (13%)
 - limitaciones:
 - no incluyen enlaces a estatutos y regulaciones
 - falta de contextualización de cada caso
 - sólo para usuarios estadounidenses



- **Patents**
 - Patentes publicadas por la USPTO
 - 18 (12,8%) M., lejos de los 140 M. de Google Patents
 - Limitaciones:
 - No posee un interfaz propio
 - Google Patents y Google Scholar no integran sus interfaces



- Búsqueda
 - Interfaz de búsqueda simple
 - Uso de filtros
 - Atajos (*Shortcuts*)

atajo	Definición	Ejemplo
author:	Busca en el campo autor	author: smith
site:	Recupera documentos alojados en un dominio web	site: harvard.edu
allintitle:	Sólo busca en el título	allintitle: science
filetype:	Identifica el tipo de formato de un documento	filetype: pdf
date range:	Localiza documentos dentro de un rango de fechas	date range: 1990-1999
~(word):	Amplia la consulta a sinónimos del término	~animal



- Problemas en la búsqueda:
 - PageRank es inadecuado para documentos científicos
 - Opacidad:
 - sólo 1000 registros
 - Redondeo de resultados
 - inconsistencias en sus filtros
 - La búsqueda por fecha devuelve resultados absurdos en función del tipo de ventana de tiempo
 - Los checks que seleccionan citas y patentes algunas veces no funcionan o muestra resultados extraños



- **Google Scholar Citations**

- Nace en Julio de 2011 y en noviembre se hace abierto
- permite el diseño de páginas web personales a cualquier investigador a partir de la bibliografía incluida en *Google Scholar*
- Indicadores bibliométricos (h-index, i10, citations)



Google scholar



Theodor Adorno

Sociology, Frankfurt School

sociology

No verified email

[Homepage](#)

Citation indices

	All	Since 2009
Citations	95778	33677
h-index	116	71
i10-index	383	280

Citations to my articles



Show: 20 1-20 [Next >](#)

Title / Author

Cited by

Year

The authoritarian personality TW Adorno, E Frenkel-Brunswik, DJ Levinson New York	11124	1950
Dialectic of enlightenment: Philosophical fragments M Horkheimer, TW Adorno, GS Noerr Stanford Univ Pr	9415 *	2002
Kulturindustrie. Aufklärung als Massenbetrug M Horkheimer, TW Adorno Dialektik der Aufklärung. Philosophische Fragmente, Frankfurt 128, 176	6379 *	1969
Dialektik der Aufklärung M Horkheimer, TW Adorno S. Fischer	5862 *	1969
Minima moralia: reflexões a partir da vida danificada TW Adorno Ática	3811 *	1992
Minima moralia: reflecții dintr-o viață mutilată TW Adorno Grupul Ed. Art	3807 *	2007
Minima moralia: reflexiones desde la vida dañada TW Adorno Akal Ediciones Sa	3806	2004
Minima moralia: reflektioner fra det beskadigede liv TW Adorno Gyldendal	3802 *	2003
Dialética do esclarecimento		

Google scholar

Search Authors

[Get my own profile](#) - [Help](#)

Follow this author

1 Follower

[Follow new articles](#)
[Follow new citations](#)

Co-authors

No co-authors



- **Google Scholar Citations**

- Ventajas

- las actualizaciones de perfiles se realiza de forma inmediata
 - el autor directamente quien selecciona sus propias publicaciones evitando ambigüedad

- Desventajas

- Manipulación de perfiles (Delgado Lopez-Cozar, et al., 2012).
 - la información de los perfiles no está normalizada ni estructurada



- Colonización

- Ocupado por oleadas (Ortega, 2015)
- Distribución sesgada
 - Hegemonía de brasileños
 - Preponderancia de informáticos y ciencias afines

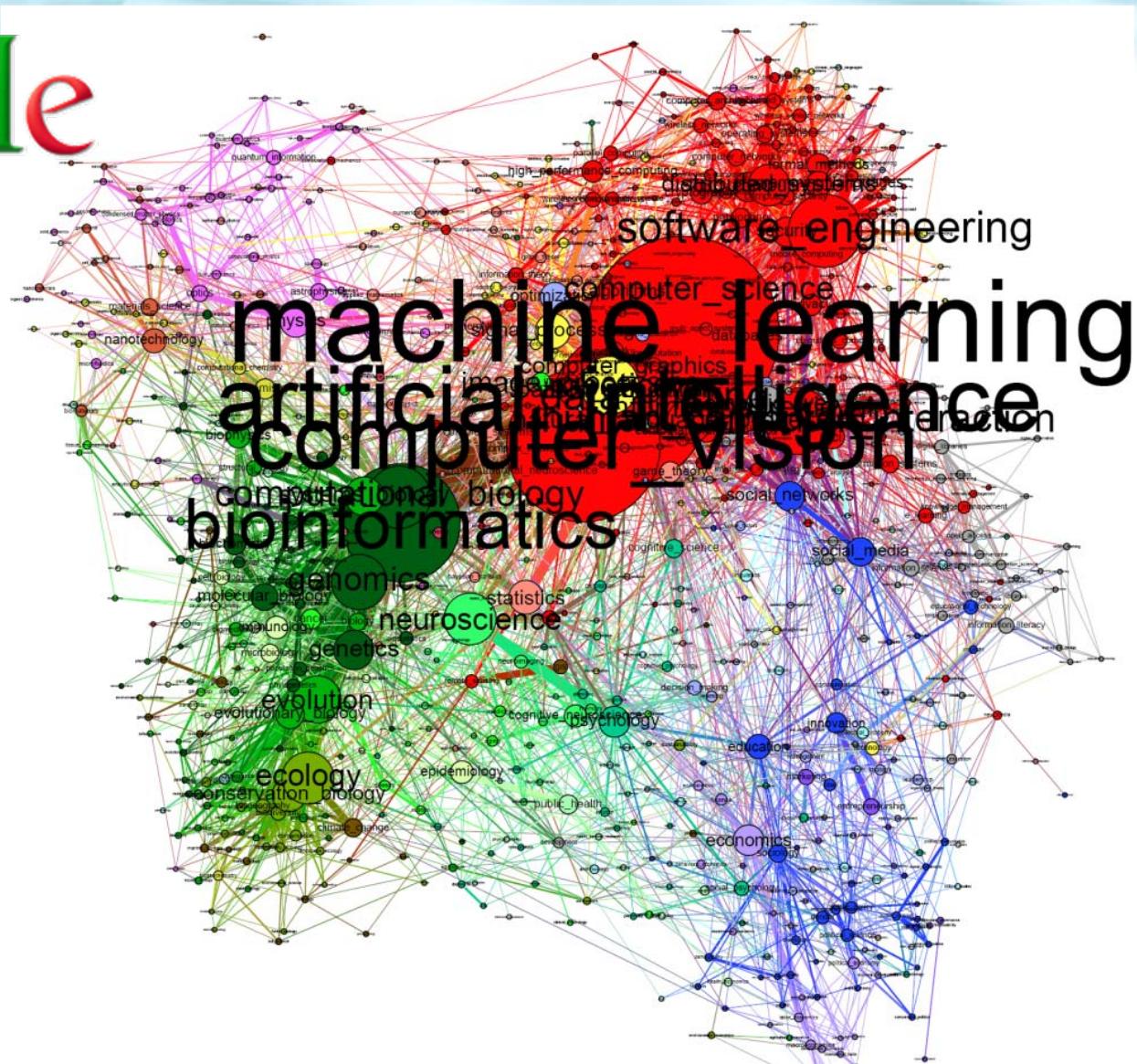
country	Profiles	%
United States	46,503	24.8
United Kingdom	12,079	6.4
Brazil	9,734	5.2
Australia	6,161	3.3
Canada	5,705	3.0
Italy	5,594	3.0
The Netherlands	3,624	1.9
Spain	3,107	1.7
Germany	3,077	1.6
Malaysia	3,005	1.6
All countries	187,301	100.0

Country	Organization	Profiles	%
Brazil	Universidade de São Paulo	3,764	1.3
Brazil	Universidade Estadual Paulista	1,636	0.6
United States	University of Michigan	1,581	0.5
Brazil	Universidade Estadual de Campinas	1,320	0.4
United States	Harvard University	1,205	0.4
United States	University of Washington	1,193	0.4
France	CNRS	1,185	0.4
United States	University of Maryland	1,146	0.4
United States	University of Minnesota	1,117	0.4
United States	University of Illinois at Urbana-Champaign	1,044	0.4
	All organizations	296,205	100.0



The Google Scholar logo consists of the word "scholar" in a smaller, semi-transparent font positioned below the main "Google" text.

label	Frequency
machine learning	4,661
artificial intelligence	4,176
computer vision	3,219
bioinformatics	2,827
data mining	2,060
software engineering	1,716
ecology	1,714
image processing	1,662
robotics	1,621
neuroscience	1,620
All labels	521,397





- **Google Scholar Metrics**

- Aparece en 2012
- seleccionar y evaluar revistas científicas a partir de las citas que reciben los artículos publicados en ellas
- Ordenado por idiomas!
- Sólo el listado en inglés está agrupado por materias
- dos indicadores propios: h5-index y h5-median
- Incluye aprox. 30.000 títulos de revistas



- **Google Scholar Metrics**

- Críticas:

- clasifica las revistas por idiomas
 - mezcla revistas con repositorios (i.e. Arxiv)
 - Sólo muestra los artículos que contribuyen al índice h
 - h5-median es una medida inútil



- Conclusiones
 - Gratuito: un 25% open access
 - Exhaustivo: cualquier material académico (tesis, informes, libros, etc.)
 - Actual: rastreo constante gracias a sus potentes robots
 - Opaco: escasa información sobre su cobertura, funcionamiento y ranking
 - Manipulable: falsos artículos en repositorios llenos de citas intencionadas
 - Pobre integración de servicios: Citations, Metrics, etc.



- Desarrollado en 2009 por Microsoft Research Asia
- A partir de dos soluciones:
 - Window Live Academic (2006): recopilación de fuentes secundarias
 - Libra (2007): autores, revistas, organizaciones, etc.



- Funcionamiento:
 - Buscador vertical multi-objeto
 - rastreo de documentos científicos en la Web
 - Identificación y extracción de los diferentes objetos.
 - Agregación de objetos al esquema previo
 - conexión entre los objetos (i.e. publicado por, escrito por, etc.)
 - Ranking: PopRank, una variación del PageRank
 - ordena los resultados teniendo en cuenta los objetos del documento



- Cobertura:
 - 39,8 M. de documentos
 - Perfiles (Window Azure, 2013)
 - 19,8 M. de autores
 - 23.000 organizaciones
 - 22.000 revistas
 - 5.200 congresos
 - 58.000 palabras clave



- Cobertura por fuentes:

Sources	Metadata	%	full-texts	%
CrossRef	40,000,000	94.23		
IEEE	2,300,000	5.42	1,800,000	12.12
ACM	80,000	.19	220,000	1.48
AGU	68,000	.16	111,000	.75
Elsevier			7,000,000	47.14
Springer			4,950,000	33.34
Arxiv.org			680,000	4.58
BioMed Central			88,000	.59
Total	42,448,000	100	14,849,000	100



- Cobertura por áreas:

Categories	Sub-categories	Papers	%	Authors	%	Journals	%
Medicine	23	12,055,802	23.53	6,011,156	21.20	5,770	28.00
Multidisciplinary	1	9,682,253	18.90	5,512,004	19.44	358	1.74
Physics	11	5,011,189	9.78	1,854,981	6.54	738	3.58
Chemistry	18	4,419,051	8.63	2,838,421	10.01	856	4.15
Biology	12	4,134,882	8.07	3,102,724	10.94	2,273	11.03
Engineering	16	3,726,867	7.27	2,568,237	9.06	1,517	7.36
Computer Science	24	3,545,328	6.92	1,600,687	5.64	1,361	6.60
Social Science	14	1,898,433	3.71	1,048,870	3.70	2,246	10.90
Arts & Humanities	14	1,373,891	2.68	538,198	1.90	1,864	9.05
Geosciences	13	1,307,164	2.55	613,753	2.16	521	2.53
Mathematics	12	1,207,282	2.36	401,751	1.42	627	3.04
Economics & Business	17	1,019,032	1.99	512,850	1.81	1,428	6.93
Material Science	13	913,525	1.78	805,963	2.84	363	1.76
Agriculture Science	8	478,967	.93	445,221	1.57	326	1.58
Environmental Sciences	10	461,486	.90	504,324	1.78	358	1.74
Total		51,235,152	100	28359140	100	20606	100



- Búsqueda
 - sólo opera sobre documentos y autores
 - no permite la búsqueda de organizaciones o revistas
 - Favorece la creación de sintaxis de búsqueda
 - navegación a través de entidades (browsing)
 - serios problemas de actualización (Abril 2013)



- Estructura
 - Autores
 - Publicaciones
 - Revistas-Congresos
 - Palabras clave
 - Organizaciones
 - existe una dimensión temática



- Autores
 - Creados a partir de los autores que aparecen en los metadatos de CrossRef
 - Duplicados: 11% del total



- Artículos
 - Biblioteca digital
 - Agrupa artículos con diferentes versiones



- Revistas-Congresos
 - Sólo hay congresos en el área de Informática y Ingeniería
 - 22 mil títulos de revista
 - 2.361 (10,7%) no contienen artículos
 - 3.819 (17,3%) no son citadas



- Revistas-Congresos por áreas:

Categories	Conferences	Conferences	Journals	Journals %	Total	Total %
Medicine	2	0.04	5,770	28.00	5,772	22.75
Computer Science	3,523	73.98	1,361	6.60	4,884	19.25
Engineering	1,235	25.93	1,517	7.36	2,752	10.85
Biology	0	0.00	2,273	11.03	2,273	8.96
Social Science	0	0.00	2,246	10.90	2,246	8.85
Arts & Humanities	0	0.00	1,864	9.05	1,864	7.35
Economics & Business	0	0.00	1,428	6.93	1,428	5.63
Chemistry	0	0.00	856	4.15	856	3.37
Physics	0	0.00	738	3.58	738	2.91
Mathematics	0	0.00	627	3.04	627	2.47
Geosciences	0	0.00	521	2.53	521	2.05
Material Science	0	0.00	363	1.76	363	1.43
Environmental Sciences	1	0.02	358	1.74	359	1.42
Multidisciplinary	1	0.02	358	1.74	359	1.42
Agriculture Science	0	0.00	326	1.58	326	1.29
Total	4,762	100	20,606	100	25,368	100



- Revistas-Congresos por fuentes:

Sources	Journals	%
Elsevier (ScienceDirect)	2,140	11.54
Springer	1,586	8.56
Wiley	1,377	7.43
Taylor & Francis	1,269	6.85
SAGE	547	2.95
Bentham Science	282	1.52
Emerald Insight	275	1.48
IEEE Xplore Digital Library	273	1.47
Inderscience	264	1.42
J-STAGE	260	1.40
All sources	18,537	



- Palabras clave
 - Asignados por el sistema, pero no se sabe como
 - 58 mil palabras clave

Categories	Keywords	Keywords %
Multidisciplinary	45,977	10.60
Engineering	39,473	9.10
Computer Science	39,010	8.99
Medicine	37,729	8.69
Biology	35,227	8.12
Physics	33,192	7.65
Chemistry	32,141	7.41
Social Science	28,348	6.53
Mathematics	26,736	6.16
Geosciences	24,902	5.74
Economics & Business	24,365	5.61
Environmental Sciences	24,033	5.54
Agriculture Science	20,380	4.70
Material Science	20,565	4.74
Arts & Humanities	1,864	0.43
All keywords	433,942	100.00



- Organizaciones

Country	Organization	Papers	Citations	Authors
China	Chinese Academy of Sciences	492,521	2,102,935	43,586
Russia	Russian Academy of Sciences	298,677	606,134	43,482
United States	Harvard University	598,801	9,336,939	42,435
United States	Stanford University	463,789	6,459,534	34,695
Brazil	University of São Paulo	186,362	747,557	32,136
United States	University of California Los Angeles	400,763	4,621,764	30,999
India	ERNET India	234,491	805,771	30,089
United States	University of California Berkeley	426,820	4,586,428	29,677
United States	University of Michigan	356,457	3,918,723	29,013
Canada	University of Toronto	329,250	3,545,927	28,835
All organizations		67,952,682	535,873,602	7,532,038



- Organizaciones
 - Doble asignación de publicaciones en el caso de traslado de autores





• Comparativa entre organizaciones



Keywords related to

New York University
Cognitive Process Confidence Interval
Decision Making Episodic Memory Fear
Conditioning Frontal Lobe Intraclass
Correlation Coefficient Intraclass
Correlation Line Drawings Mental Disorder
Negative Symptom Prefrontal Cortex
Procedural Justice Recognition Memory Self
Efficacy Social Identity Social Phobia **Statistical**
Model Statistical Significance Temporal Lobe
Unconditional Stimulus White Matter **Working Memory**
Young Adult posttraumatic stress disorder

Both Organizations
Children and Adolescents Clinical Trial **Cognitive**
Deficit College Students Control Group Control
Subjects Cortisol Executive Function **Gender**
Difference High Risk Hypothesis Test
Indexation **Individual Difference Low**
Income Obsessive Compulsive Disorder
Rating Scale Risk Factors **Sample Size Side**
Effect Social Behavior Social Cognition Social
Support Substance Use **Treatment Outcome**
Young Children

University of North Carolina Chapel Hill
Aggressive Behavior Antipsychotic Drug Atypical
Antipsychotics **Bipolar Disorder Blood Pressure**
Child Care **Child Development Covariance**
Structure Eating Disorder Emotion Regulation
Empirical Evidence Exploratory Factor Analysis
Factor Analysis Negative Affect Physical Health
Positive Emotion Positive Psychology
Prospective Study Serotonin Sex Difference
Structural Equation Model Treatment Effect
Twin Study anorexia nervosa bulimia nervosa



- **Visual Explorer**

- Co-author graph: grafo de los principales coautores de un autor (hasta 30)
- Co-author path: la ruta que separa a dos autores a través de coautorías
- Citation graph: autores que citan a un autor
- Genealogy graph:
 - ya desaparecido
 - relación entre directores de tesis y doctorandos.



Microsoft Academic Search Beta

Academic > Author > Jose Luis Ortega Priego >

jose luis ortega

Result

- Jose Luis Ortega P... Consejo Superior ...
- Jose Luis Ortega New Mexico State ...
- Jose Luis Ortega (...)
- Jose Luis Diaz-Ort... Instituto Naciona...
- Jose Luis Gallego ... Instituto Valenci...
- JOSE LUIS GODINEZ... Universidad Natio...
- Jose Luis Lozano O...
- Jose Luis Ortega-M... Universidad de Gr...
- Jose Luis Ortega-R...
- Jose Luis Ortega-S...
- Jose Luis Ruiz Ort...
- Jose Luis Velazque...
- Jose-Luis Ortega-R... University of Oxf...
- Por Jose Luis Orte...

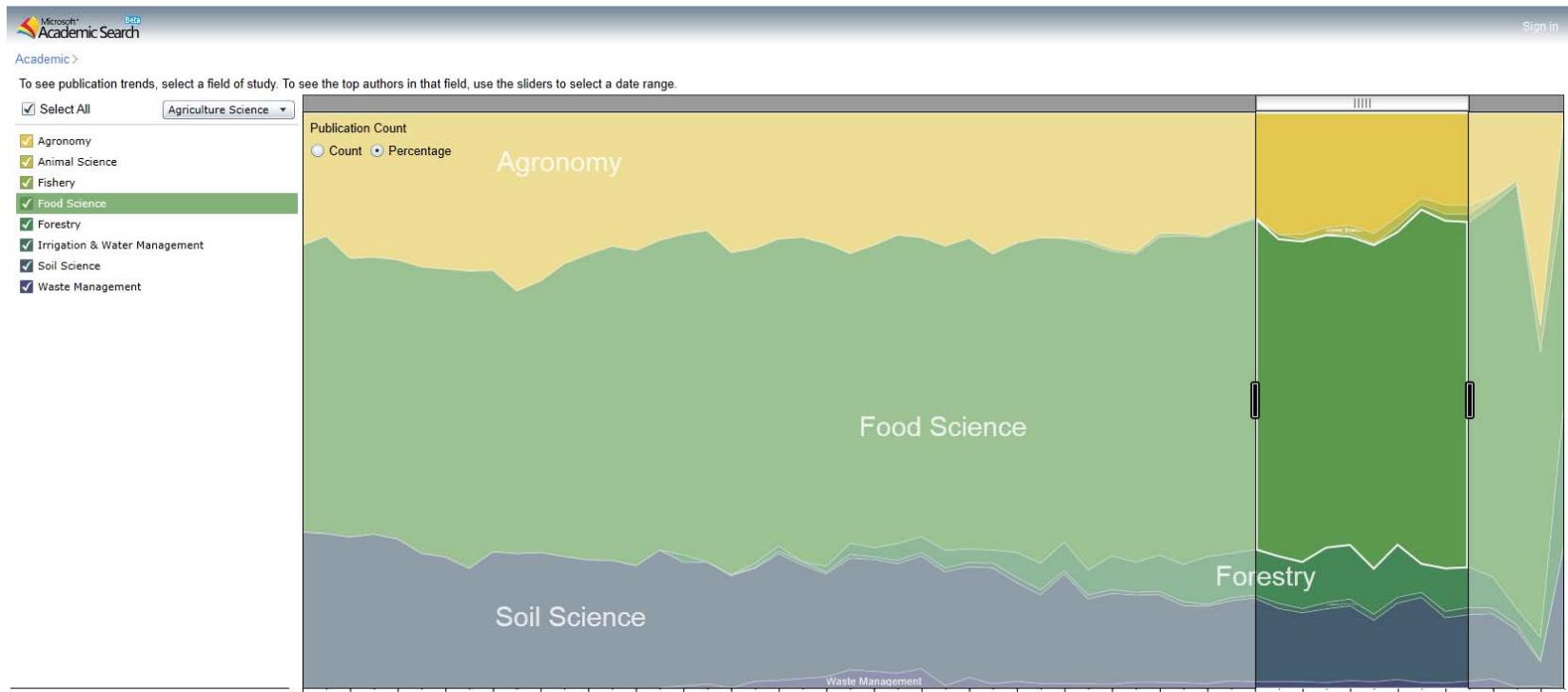
Co-author Graph Co-author Path Citation Graph

The visualization shows a network of academic authors connected by lines representing co-authorships. The central node is Jose Luis Ortega Priego. Other prominent nodes include Isidro F. Aguilera, Ana Maria Utrilla, and Daniel Madrid. Lines connect the central node to various other authors, such as Vicky, Marilyn Fernandez, and Encarnacion Ramos Hidalgo. The size of the nodes varies, likely indicating the number of publications or citations. The background is light blue, and the nodes are represented by blue circles with profile icons.



Microsoft®
Academic Search Beta

• Publication trends



Top Authors in Food Science (2000 - 2009) [More >](#)

Soottawat Benjakul
Prince of Songkla University

David Julian McClements
University of Massachusetts

Da-Wen Sun
University College Dublin

Dietrich Knorr
Technical University of Berlin

Fereidoon Shahidi
Memorial University of Newfoundland



- Conclusiones
 - Gratuito: biblioteca virtual
 - Estructurado: distintos niveles de agrupación
 - Visualización y comparativas: información elaborada
 - Desactualizado: desde abril de 2013
 - Mala gestión de duplicados:
 - Autores repetidos
 - Artículos no agrupados



Conclusiones generales

- No existe una definición clara de buscador académico
- Graves deficiencias en la recuperación
 - Búsquedas avanzadas pobres
- Nuevas perspectivas a la evaluación científica
 - Comparativas
 - Visualización
 - Entidades (autores, organismos, revistas, etc.)
- Jóvenes con gran potencial



Bibliografía

- Agullo, I. F. (2012). Is Google Scholar useful for bibliometrics? A webometric analysis. *Scientometrics*, 91(2), 343-351
- Delgado López-Cózar, E., Robinson-García, N., & Torres-Salinas, D. (2014). The Google Scholar Experiment: how to index false papers and manipulate bibliometric indicators. *Journal of the Association for Information Science and Technology*, 65(3), 446-454.
- Khabsa, M., & Giles, C. L. (2014). The Number of Scholarly Documents on the Public Web. *PloS one*, 9(5), e93949
- Orduña-Malea, E., Ayllón Millán, J. M., Martín-Martín, A., & Delgado López-Cózar, E. (2014). How many academic documents are visible and freely available on the Web?
- Ortega, J. L. (2015). How is an academic social site populated? A demographic study of Google Scholar Citations population. *Scientometrics*
- Microsoft (2014). Microsoft Academic | Window Azure Marketplace. Retrieved March 22, 2014, from <http://datamarket.azure.com/dataset/mrc/microsoftacademic>
- Ortega, J. L. and Agullo, I. F. (2012), Science is all in the eye of the beholder: Keyword maps in Google scholar citations. *Journal of the American Society for Information Science and Technology*, 63: 2370–2377
- Ortega, J. L. and Agullo, I. F. (2014). Microsoft academic search and Google scholar citations: Comparative analysis of author profiles. *Journal of the Association for Information Science and Technology*, 65(6), 1149-1156.



Gracias!!

- Si quieres saber más...



- Ortega, J. L. (2014). Academic Search Engines: A quantitative outlook. Cambridge, UK: Chandos Publishing (Elsevier) ISBN:9781843347910