

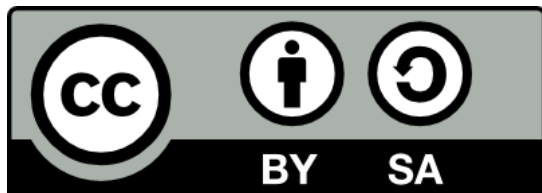
A Cloud-Based Infrastructure for Community-Driven Data Mining on Research Outputs

Robert Jäschke^{1,3} & Andreas Hotho^{2,3}

¹The University of Sheffield, UK

²The University of Würzburg, Germany

³L3S Research Center Hannover, Germany

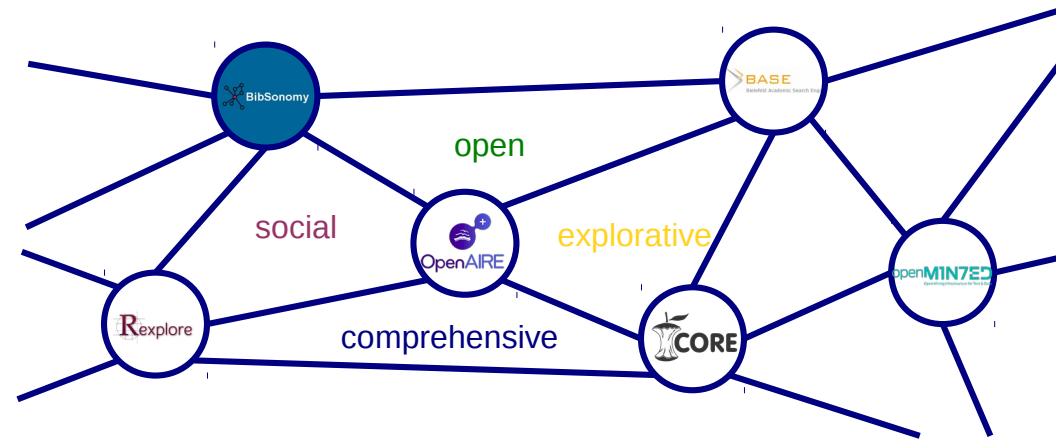


Information
School.



Agenda

1.future (what we need)

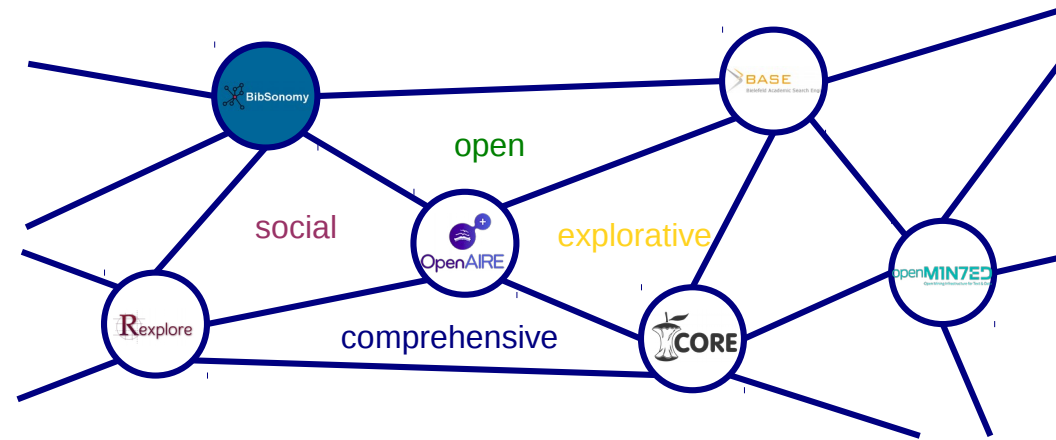


2.past & present (what we have)



Agenda

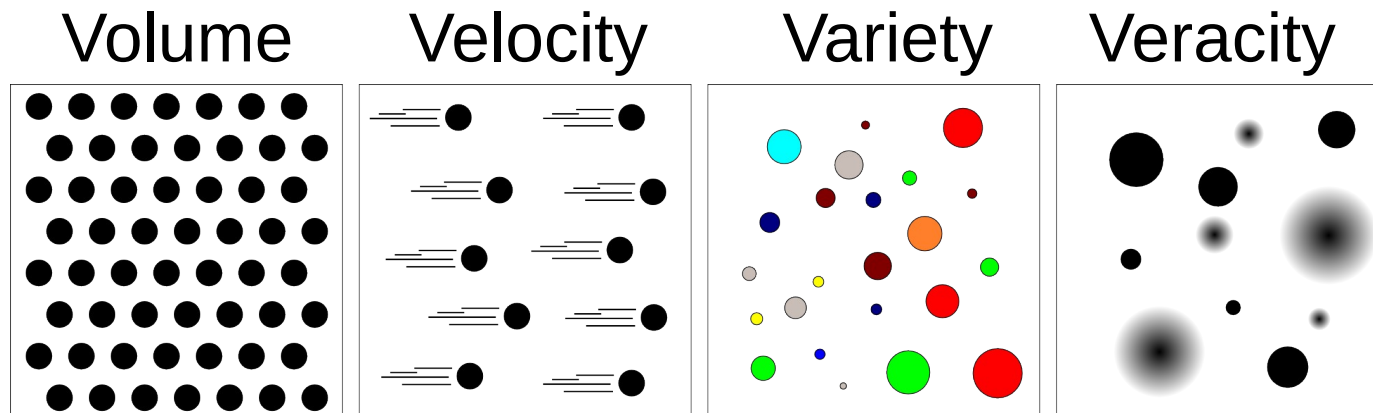
1.future (what we need)



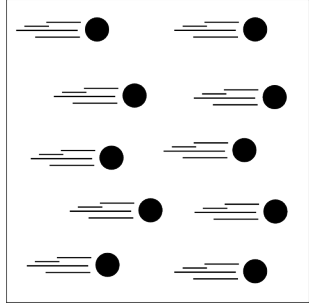
2.past & present (what we have)



Research Outputs are Big Data

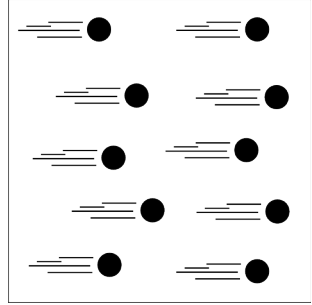


Velocity



*In every 24-hour period approximately 20,000,000 words of technical information are being recorded. A reader capable of reading 1,000 words per minute would **require 1.5 months**, reading 8 hours every day, **to get through 1 day's technical output**, and at the end of that period, he would have fallen 5.5 years behind in his reading!*

Velocity



In order to emphasize the problem, it is worthwhile to cite some figures. The rate at which technical documents are produced at the present time is estimated to be well over 500,000 per year. In every 24-hour period approximately 20,000,000 words of technical information are being recorded. A reader capable of reading 1,000 words per minute would require $1\frac{1}{2}$ months, reading 8 hours every day, to get through 1 day's technical output, and at the end of that period, he would have fallen $5\frac{1}{2}$ years behind in his reading! Even in attempting to read the portion of the literature in a single subject field such as chemistry, he would find himself falling behind an estimated 850,000 pages per year. This production rate of scientific information will undoubtedly increase as countries such as China and India begin to produce technical work commensurate with their size.

Velocity

REPORT NO. RSIC-510

**METHODS FOR SATISFYING THE NEEDS
OF THE SCIENTIST AND THE ENGINEER
FOR SCIENTIFIC AND TECHNICAL INFORMATION**

by

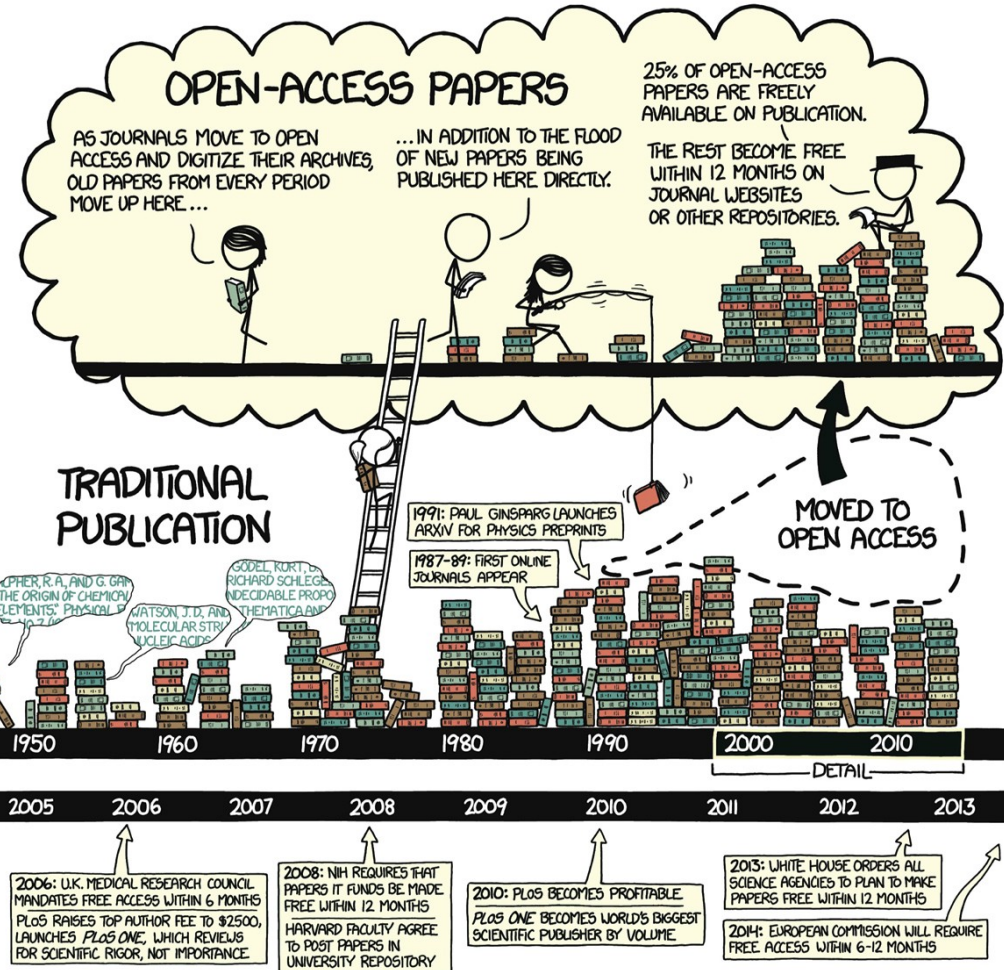
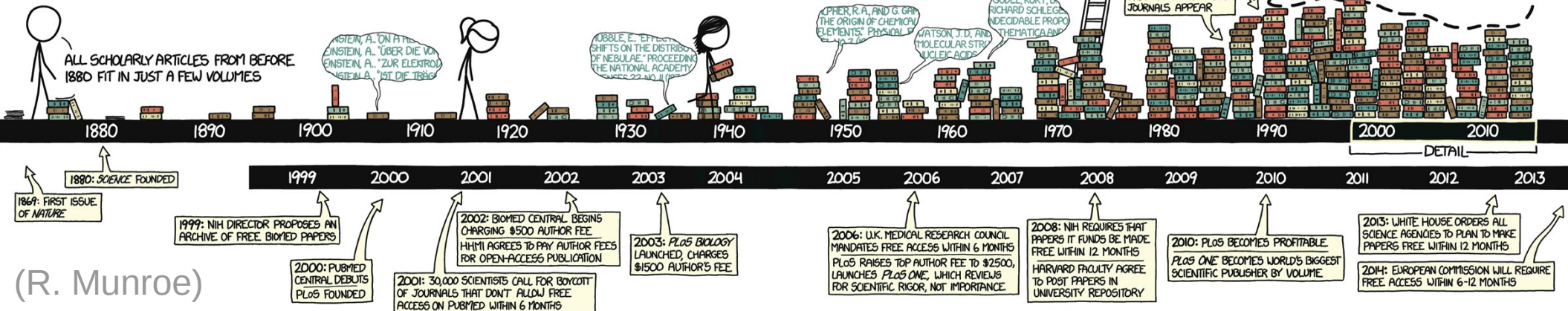
Hubert Murray, Jr.

January 1966

CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION			
Hardcopy	Microfiche		
\$ 1.00	\$ 0.50	19 pp	as
ARCHIVE COPY			

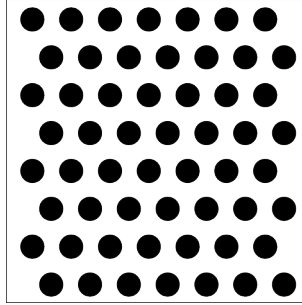
Code 1

A 7x7 grid of black dots. The dots are arranged in 7 rows and 7 columns. The dots on the main diagonal (from top-left to bottom-right) are missing, leaving 42 dots in total.



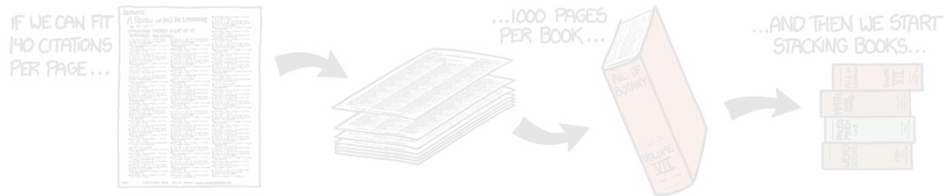
8

Volume



HOW MUCH SCIENCE IS THERE?

SCIENTIFIC PUBLISHING HAS BEEN ACCELERATING—A NEW PAPER IS NOW PUBLISHED ROUGHLY EVERY 20 SECONDS. LET'S IMAGINE A BIBLIOGRAPHY LISTING *EVERY* SCHOLARLY PAPER EVER WRITTEN. HOW LONG WOULD IT BE?

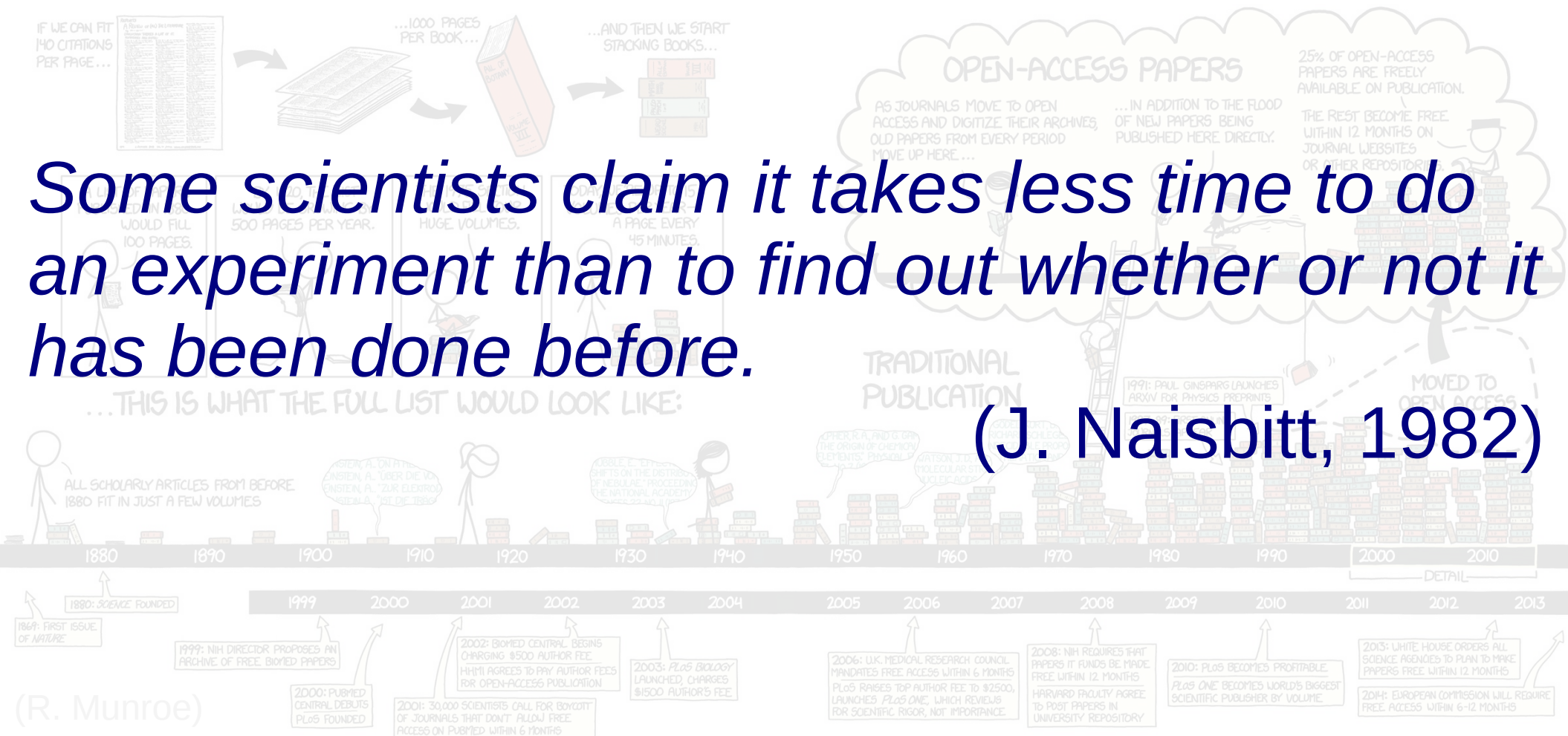


HOW OPEN IS IT?

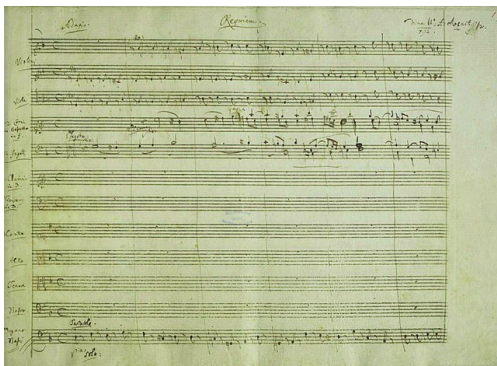
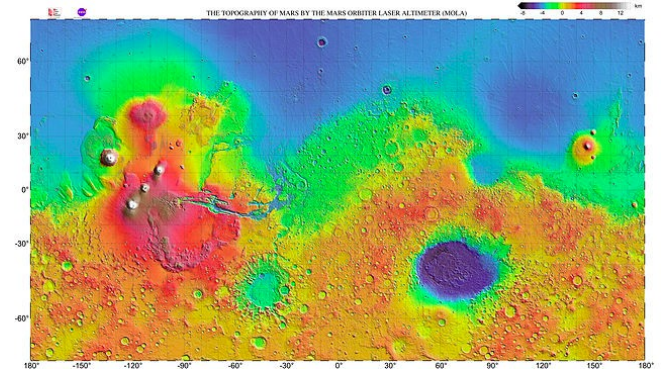
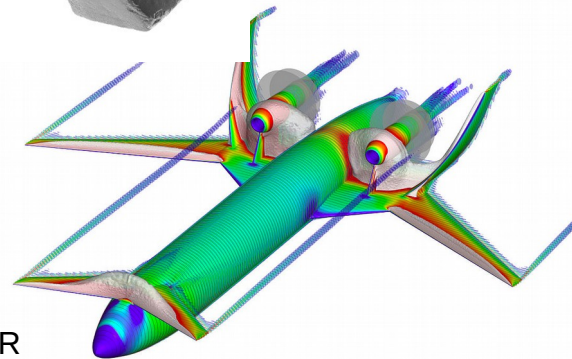
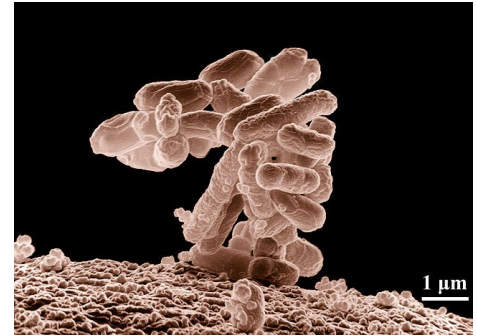
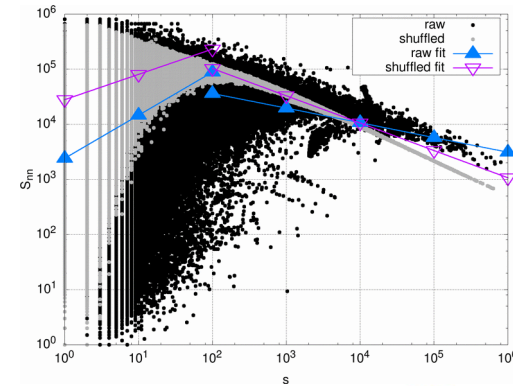
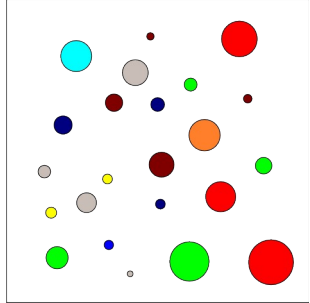
SINCE THE ADVENT OF THE WEB, MUCH OF SCIENTIFIC PUBLISHING HAS BEEN MOVING TO *OPEN ACCESS*. ACCORDING TO SCIENCE-METRIX, OPEN ACCESS REACHED A "TIPPING POINT" AROUND 2011: MORE THAN 50% OF NEW RESEARCH IS NOW MADE AVAILABLE FREE ONLINE.

Some scientists claim it takes less time to do an experiment than to find out whether or not it has been done before.

(J. Naisbitt, 1982)



Variety



DLR

Sun Jul 27 17:25:36 +0000 2014	493447121828212738	30198750	http://www.nature.com/n
Thu Jan 16 02:10:44 +0000 2014	423638418158272512	30198750	http://elemente.wordpress
Sun Sep 07 03:39:18 +0000 2014	508459468288720896	30198750	https://sites.google.co
Fri Mar 14 16:08:23 +0000 2014	444505330060636160	30198750	http://raetschlab.org/s
Thu Sep 25 22:31:16 +0000 2014	515267316201054208	30198750	http://people.tuebingen
Wed Oct 01 20:39:41 +0000 2014	517413562357403648	30198750	http://projectreporter.
Wed Jun 04 23:57:22 +0000 2014	474339158035816448	30198750	http://hitseq2014.sched
Thu Jul 10 00:28:03 +0000 2014	487030453451251712	30198750	http://hitseq.org/
Mon Jan 06 21:00:43 +0000 2014	420298907604627456	30198750	http://www.nytimes.com/
Fri Oct 17 03:03:26 +0000 2014	522945956661059584	30198750	https://www.facebook.co
Wed Oct 01 20:39:41 +0000 2014	517413562357403648	30198750	http://projectreporter.
Sun Dec 14 11:20:22 +0000 2014	544089511278309377	30198750	http://www.biomedcentra
Sat Oct 18 18:04:34 +0000 2014	523535121753063424	30198750	https://github.com/ga4g
Tue Sep 02 13:11:57 +0000 2014	506791641718722560	30198750	http://www.msccc.org/mo
Tue Dec 23 16:00:44 +0000 2014	547421557669507072	30198750	http://www.fredericksbu
Sun Jan 12 08:01:07 +0000 2014	42227745154893824	30198750	http://www.slideshare.n
Wed Oct 01 20:43:29 +0000 2014	517414522387431425	30198750	http://projectreporter.
Sat Oct 18 18:09:11 +0000 2014	523536283596251138	30198750	https://github.com/ga4g
Thu Oct 23 01:01:23 +0000 2014	525089569407258625	30198750	http://ml4chg.org/
Wed Oct 01 20:41:57 +0000 2014	517414133940379648	30198750	http://projectreporter.
Thu Nov 20 00:35:59 +0000 2014	535230035456978944	30198750	http://www.nytimes.com/
Tue Dec 09 14:17:26 +0000 2014	542322130105688064	30198750	http://ml4chg.org/
Tue Mar 11 18:58:53 +0000 2014	443461071505215488	30198750	http://nyas.org/ML2014
Tue Nov 11 04:15:45 +0000 2014	532023852462010368	30198750	http://m.imdb.com/title

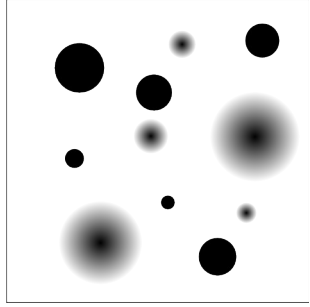
```

for row in reversed(rows):
    url = self.urls[row]
    if self.url_is_publisher(url):
        # a true positive found
        tp_rate.append(tp_rate[-1] + 1)
        fp_rate.append(fp_rate[-1])
        tp_count += 1
    else:
        # a false negative found
        tp_rate.append(tp_rate[-1])
        fp_rate.append(fp_rate[-1] + 1)
        fp_count += 1

# remove first (artificial) element
tp_rate.pop(0)
fp_rate.pop(0)

return np.array(fp_rate)/float(fp_count),
    
```

Veracity

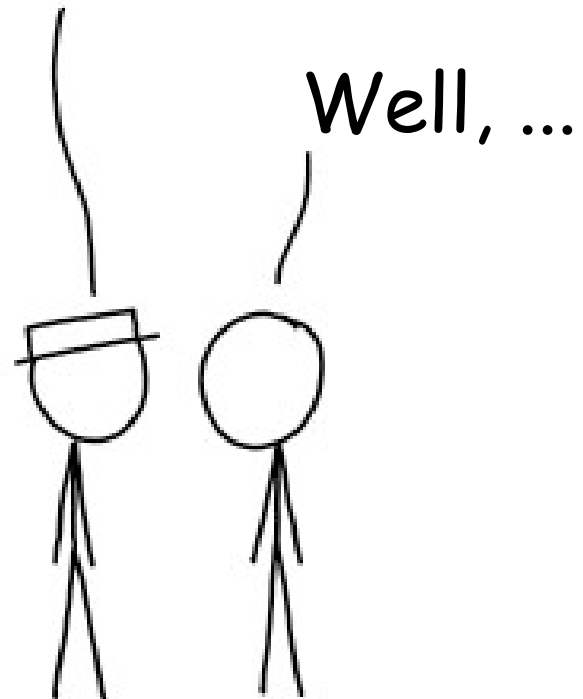


REFERENCES

16. NEED CITATION FROM TELECOM-CITIES
17. PBS. (n.d.). *Forensic anthropology*. Retrieved September 24, 2005, from PBS Web site:
<http://www.pbs.org/opb/historydetectives/techniques/forensic.html>
18. Multidisciplinary Center for Earthquake Engineering Research (2005) Ibid.
19. Washington State Department of Transportation. (1940) "*Gallopig Gertie*" collapses November 7, 1940.
Retrieved September 30, 2005, from Washington State Department of Transportation Web site:
<http://www.wsdot.wa.gov/TNBhistory/Connections/connections3.htm>
20. Institute for the Future. NEED CITATION FROM TEN YEAR FORECAST?

State of the Art

How do we tackle this issue?



(R. Munroe, XKCD)

State of the Art

About 74,700 results (0.04 sec)

Recommender systems

P. Resnick, H.R. Varian - Communications of the ACM, 1997 - di.ac
Recommender systems assist and augment this natural social pri
recommender system people provide recommendations as inputs,
aggregates and directs to appropriate recipients. In some cases t
Cited by 3805 Related articles All 6 versions Cite Save Mor

Toward the next generation of **recommender systems**
of-the-art and possible extensions

G. Adomavicius, A. Tuzhilin - IEEE transactions on knowledge ...
Abstract—This paper presents an overview of the field of **recomm**
describes the current generation of recommendation methods that
the following three main categories: content-based, collaborative,
Cited by 7273 Related articles All 30 versions Cite Saved

[PDF] Matrix factorization techniques for **recommender systems**
Y. Koren, R. Bell, C. Volinsky - Computer, 2009 - datajobs.com
Figure 2 illustrates this idea for a simplified example in two dimen
hypothetical dimensions characterized as female-versus male-orie
escapist. The figure shows where several well-known movies and
Cited by 2667 Related articles All 44 versions Cite Saved

[BOOK] Introduction to **recommender systems** handbook
R. Ricci, L. Rokach, B. Shapira - 2011 - Springer
Abstract **Recommender Systems (RS)** are software tools and t
suggestions for items to be of use to a user. In this introduction c
basic RS ideas and concepts. Our main goal is to delineate, in a
Cited by 1877 Related articles All 25 versions Cite Save

Evaluating collaborative filtering **recommender systems**
J.L. Herlocker, J.A. Konstan, L.G. Tene - ... on Information Syste
Abstract **Recommender systems** have been evaluated in many, c
this article, we review the key decisions in evaluating collaborati
systems: the user tasks being evaluated, the types of analysis an
Cited by 4319 Related articles All 22 versions Cite Saved

Hybrid **recommender systems**: Survey and experim
R. Burke - User modeling and user-adapted interaction, 2002 - Spri
Abstract **Recommender systems** represent user preferences for
items to purchase or examine. They have become fundamental ap
commerce and information access, providing suggestions that effe
Cited by 2941 Related articles All 30 versions Cite Saved

Recommender systems in e-commerce
J.B. Schafer, J. Konstan, J. Riedl - Proceedings of the 1st ACM conf
ABSTRACT **Recommender systems** are changing from novelties
sites, to serious business tools that are re-shaping the world of E-
largest commerce Web sites are already using **recommender sys**
Cited by 1639 Related articles All 21 versions Cite Save

Being accurate is not enough: how accuracy metric
systems
S.M. McNee, J. Riedl, J.A. Konstan - ... Human factors in computing
Abstract **Recommender systems** have shown great potential to h
and relevant items from within a large information space. Most res
focused on improving the accuracy of **recommender systems**. W
Cited by 587 Related articles All 7 versions Cite Save

Context-aware **recommender systems**
G. Adomavicius, A. Tuzhilin - **Recommender systems** handbook, ...
Abstract The importance of contextual information has been recog
practitioners in many disciplines, including e-commerce personal
retrieval, ubiquitous and mobile computing, data mining, marketing
Cited by 1060 Related articles All 29 versions Cite Saved

Trust in **recommender systems**
J. O'Donovan, B. Smyth - Proceedings of the 10th international con
Abstract **Recommender systems** have proven to be an important
overload problem, by providing users with more proactive and pers
services. And collaborative filtering techniques have proven to be a
Cited by 772 Related articles All 15 versions Cite Saved



Search results for recommender systems (1309923 articles were found):

First Previous 1 2 3 ... Next Last

Recommender Systems

Provided by arXiv.org e-Print Archive | 2012

By Linyuan Lu, Matus Medo, Chi Ho Yeung, Yi-Cheng Zhang, Zi-Ke Zhang and

Recommender Systems Linyuan Lu^{a,b,c}, Matus Medo, Chi Ho Yeung^d, Yi-Cheng Zhang^a,
Zhoua,b,c,e aInstitute of Information Economy, Hangzhou Normal University, Hangzhou, 310036
Physics, University of Fribourg, Fribourg, CH-1700, Switzerland cWeb Sciences Center, Universi
Technology of China, Chengdu, 610054, PR China dThe Nonlinearity and Complexity Research
Birmingham B4 7ET, United Kingdom...

Item reputation-aware recommender systems

Provided by Queensland University of Technology ePrints Archive | Publisher:

By Ahmad Abdel-Hafez, Yue Xu and Nan Tian

...This is the author's version of a work that was submitted/accepted for pub- lication in the follow
Xue, Y. & Tian, Nan (2014) Item reputation-aware **recommender systems**. In IWAS 14: Proce
on Information In- tegration and Web-based Applications & Services, ACM Digital Library, Haroi
downloaded from: http://eprints.qut.edu.au/7819/ c© Copyright 2014 ACM Notice: Changes in

Hybrid critiquing-based recommender systems

Provided by Infoscience - École polytechnique fédérale de Lausanne | Publish
Computing Machinery, New York, NY 10036-5701, United States | 2007

By Li Chen and Pearl Pu

... Hybrid Critiquing-based **Recommender Systems** Li Chen and Pearl Pu Human Computer Inte
Computer and Communication Sciences Swiss Federal Institute of Technology in Lausanne (EP
Switzerland (li.chen, pearl.pu)@epfl.ch ABSTRACT We propose a novel critiquing-based **recom**
critiquing interface that integrates the user self-motivated critiquing facility to compensate for the
critiques...

Single-class recommender systems

Provided by ePrints.FRI | 2011

By Matic Perovšek

Efficient Recommender Systems

Provided by Research Papers in Economics

By Dirk Bergemann and Deran Ozmen

... EFFICIENT **RECOMMENDER SYSTEMS** By Dirk Bergemann and Deran Ozmen June 2006 C
DISCUSSION PAPER NO. 1568 COWLES FOUNDATION FOR RESEARCH IN ECONOMICS Y
New Haven, Connecticut 06520-8281 http://owles.econ.yale.edu/ Efficient **Recommender Syst**
University New Haven, CT 06511 dirk.bergemann@yale.edu Deran Ozmen Yale University New
deran.ozmen@yale.edu April 2006...

trust recommender

Provided by EdShare

By

...likely to get the same results. Another approach to the information overload problem are **recom**
proposes an agent based **recommender** system based on the concept of trust in social networks
quantify a measure of trust relationships in the context of the social networks and then designs a
recommender system. Such a **recommender** system is originally proposed to be used as an info
auction...

seconds)

Sort by: Relevance

Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions

2005, **IEEE Transactions on Knowledge and Data Engineering**, volume 17, issue 6, pp 734-749

Gediminas Adomavicius (*University of Minnesota*), Alexander Tuzhilin

This paper presents an overview of the field of recommender systems and describes the
current generation of recommendation methods that are usually classified into the following
three main categories: content-based...

Fields of Study: collaborative filtering, context model, collaboration, ...

Cited 7635 times*

Evaluating collaborative filtering recommender systems

2004, **ACM Transactions on Information Systems**, volume 22, issue 1, pp 5-53

Jonathan L. Herlocker (*Oregon State University*), Joseph A. Konstan (*University of Minnesota*),
Loren Terveen (*University of Minnesota*), John Riedl (*University of Minnesota*)

Recommender systems have been evaluated in many, often incomparable, ways. In this article,
we review the key decisions in evaluating collaborative filtering recommender systems: the
user tasks being evaluated, the...

Fields of Study: collaborative filtering, metrics, evaluation, ...

Cited 4635 times*

Item-based collaborative filtering recommendation algorithms

2001, **World Wide Web**

Badrul M. Sarwar (*University of Minnesota*), George Karypis (*University of Minnesota*), Joseph
A. Konstan (*University of Minnesota*), John Riedl (*University of Minnesota*)

Fields of Study: dynamic data, collaborative filtering, dissemination, ...

Cited 5938 times*

Matrix Factorization Techniques for Recommender Systems

2009, **IEEE Computer**, volume 42, issue 8, pp 30-37

Yehuda Koren (*Yahoo!*), Robert M. Bell, Chris Volinsky

As the Netflix Prize competition has demonstrated, matrix factorization models are superior to
classic nearest neighbor techniques for producing product recommendations, allowing the
incorporation of additional information such as implicit...

Fields of Study: matrix decomposition, **recommender system**, computational intelligence, ...

Cited 2691 times*

Social information filtering: algorithms for automating "word of mouth"

1995, **Human Factors in Computing Systems**

Utpendra Sharanand (*Massachusetts Institute of Technology*), Pattie Maes (*Massachusetts
Institute of Technology*)

Fields of Study: word of mouth, user modeling, **recommender system**, ...

Cited 3887 times*

Factorization meets the neighborhood: a multifaceted collaborative filtering model

2008, **Knowledge Discovery and Data Mining**

Yehuda Koren (*AT&T*)

2016-09-29

13

Vision

make the state of the art accessible

The figure is a composite image illustrating a deep learning application in computer vision, specifically related to point cloud processing or autonomous driving perception.

Top Left: A 3D visualization of a green airplane model, likely generated from a point cloud or depth map, shown against a white background with coordinate axes.

Top Right: A heatmap visualization of a scene, possibly a road or terrain, showing various colored regions (red, yellow, green, blue) representing different classes or features. The x and y axes are labeled from 0 to 100.

Bottom Left: A large, dark blue diagonal watermark reading "Comprehensive".

Bottom Center: A list of URLs, each preceded by a timestamp (e.g., 2018-07-27 10:11:17 AM UTC+08:00). These URLs appear to be sources for datasets or models used in the project.

Bottom Right: A Python code snippet demonstrating operations on a NumPy array named `fp_rate`:

```
for i in reversed(rows):  
    for j in range(1, cols):  
        if not np.isnan(fp_rate[url]):  
            # valid active found  
            tp_rate[j] = max(tp_rate[j - 1] + 1,  
                             fp_rate[j] - fp_rate[j - 1])  
            fp_rate[j] = fp_rate[j - 1]  
        else:  
            # invalid active found  
            tp_rate[j] = max(tp_rate[j - 1],  
                             fp_rate[j] - fp_rate[j - 1] + 1)  
            fp_rate[j] = fp_rate[j - 1]  
        j += 1  
# remove first column (col=0)  
tp_rate.pop(0)  
fp_rate.pop(0)  
return np.array(fp_rate)/(fp_rate + 1)
```

open



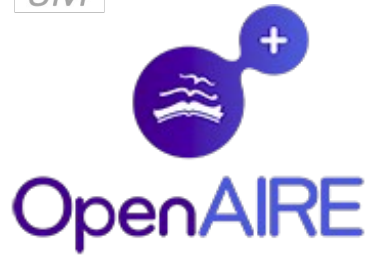
The collage features several data visualization techniques:

- Bar Charts:** A grouped bar chart at the top left shows data from 2000 to 2010. A line chart at the bottom left shows data from 1980 to 2000.
- Network Graphs:** A circular network graph at the top right shows connections between nodes like 'Yang Ty', 'Gerdard Nykum', and 'Information Regional, WWW, SA'. A hierarchical tree diagram at the bottom right shows a structure with nodes like 'Knowledge Base, AI, DA', 'Semantic Web', and 'Ontology, Ontology Mapping, OA'.
- Sankey Diagram:** A Sankey diagram in the center shows the flow of data between different categories, with a large red flow and a smaller blue flow.
- Watermark:** A large, yellow, diagonal watermark with the word 'explorative' is overlaid across the entire image.



Comprehensive

CiteSeer^x 8M



open



social

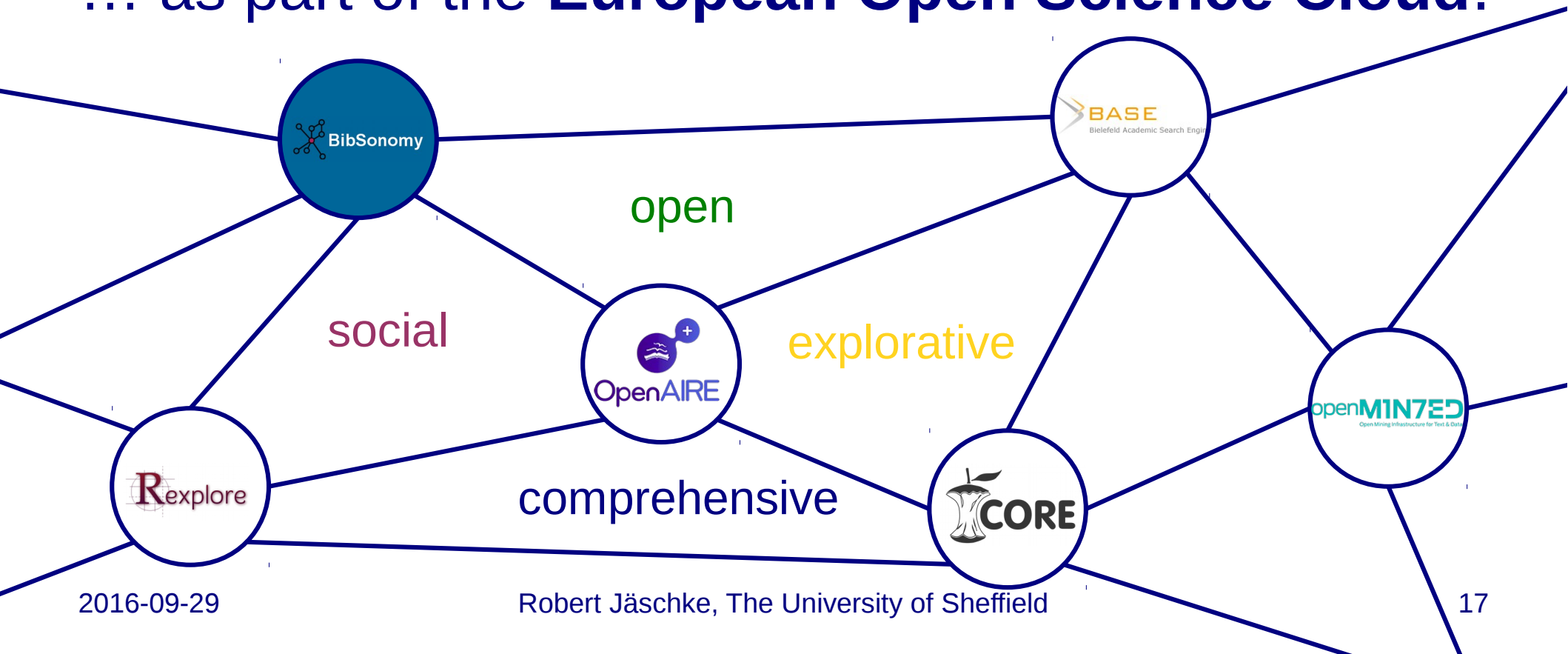


Proposal

Combine the power of these systems ...

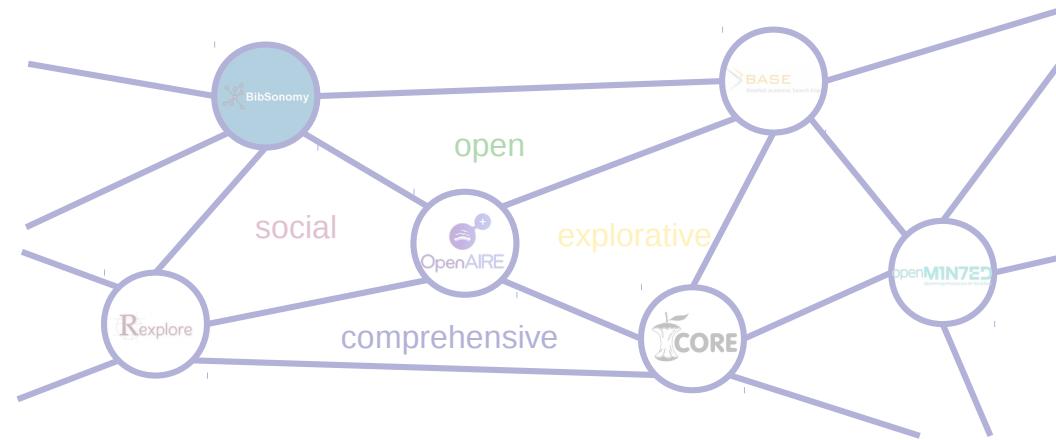
... into a federated, open platform ...

... as part of the **European Open Science Cloud**.



Agenda

1.future (what we need)



2.past & present (what we have)



bookmarks 1228



OpenCitations - Home

1



<http://opencitations.net/>

🕒 2 days ago by @jaeschke

🔖 citation data dataset open

★★★★★ (0)



publications 1082



Tropical Truth(s): The Epistemology of Metaphor and ot...

1



A. Burkhardt, and B. Nerlich (Eds.). *De Gruyter, (2010)*

🕒 5 days ago by @jaeschke

🔖 dh epistemology language metaphor text trope

★★★★★ (0)



Handbuch Open Science – Wikibooks, Sammlung freie...

1



https://de.wikibooks.org/wiki/Handbuch_Open_Science

🕒 23 days ago by @jaeschke

🔖 handbuch open science wikibooks

★★★★★ (0)



Why does deep and cheap learning work so well?

3



H. Lin, and M. Tegmark. (2016) *cite arxiv:1608.08225* Comment: 14 pages, 3 figs.

🕒 6 days ago by @jaeschke

🔖 deep learning machine network neural

★★★★★ (0)



Comparison of research networking tools and research...

1



https://en.wikipedia.org/wiki/Comparison_of_research_...

🕒 2 months ago by @jaeschke

🔖 gaw network profile researcher

★★★★★ (0)



Web Archiving: Issues and Methods

1



J. Masanès. *page 1--53. Springer, Berlin/Heidelberg, (2006)*

🕒 12 days ago by @jaeschke

🔖 archive web

★★★★★ (0)



Hermitage: Testing the "I" in ACID — Martin Kleppman...

1



<https://martin.kleppmann.com/2014/11/25/hermitage-tes...>

🕒 2 months ago by @jaeschke

🔖 acid base database isolation nosql sql

★★★★★ (0)



Determining the Optimal Session Interval for Transactio...

2



S. Wakeling, and P. Clough. *page 703--708. Springer, (2016)*

🕒 14 days ago by @jaeschke

🔖 catalogue digital dl interval library log session

★★★★★ (0)



@jaeschke ↗

Robert Jäschke



You must be signed in to add this user to your friend list.

sign in

discussion

show posts recently discussed by jaeschke

concepts

similar users

(more...)

@macdet @schmitz @flint63 @dolefulrabbit
@stumme @hotho @beate @dbenz @sdo
@brightbyte

shared groups

tags

alexandria analysis archive

bibsonomy bibsonomynews

blog bookmarking citation collaborative

collective community computing concept data

dataset fca folksonomy fotw geo

graph information intelligence java library

map mining myown network ontology

open paper programming publication

recommender research science

scientometrics search semantic sna

social survey tag tagging toread twitter

visualization web wikipedia

[🏠](#) / [myBibSonomy](#) / The Social Bookmark an...


The Social Bookmark and Publication Management System BibSonomy

Dominik Benz, Andreas Hotho, Robert Jäschke, Beate Krause, Folke Mitzlaff, Christoph Schmitz, and Gerd Stumme.
The VLDB Journal 19(6):849--875 (December 2010)



Abstract

Social resource sharing systems are central elements of the Web 2.0 and use the same kind of lightweight knowledge representation, called folksonomy. Their large user communities and ever-growing networks of user-generated content have made them an attractive object of investigation for researchers from different disciplines like Social Network Ana... (more)

Links and resources

DOI: 10.1007/s00778-010-0208-4

URL: <http://www.kde.cs.uni-kassel.de/pub/pdf/benz2010social.pdf>

additional links:

+ add URL

OpenURL: <https://getinfo.de/app/openurl/search?date=2010&id=doi%3A10.1007%...>

BibTeX key: benz2010social

Internal link:

search on:

Comments and Reviews ★★★★★ (1)

anonymous has written a comment or review. [Join the discussion!](#)

Tags

2010 bibsonomy bookmark collaborative folksonomy kde management myown publication sent:hotho system tagging top vldb web-observatory

Cite this publication



Meta data

Last update a year and 3 months ago
 Created 4 years and 10 months ago

edit tags

2010 bibsonomy bookmark coll

community

In collection of:

- @dmir
- @kde-alumni
- @kde-alumni
- @bibsonomy-ruby
- @michael-t-weber
- @thoni
- @asxc0312
- @schmidt2
- @schmitz
- @kibanov
- @l3s
- @olya
- @folke
- @thoni
- @nosebrain
- @victorinostar
- @iuse-bilal

www.bibsonomy.org

History



BibSonomy::home

search [home](#) :: [tags](#) :: [popular](#)username: password: [help](#)[login](#) :: [register](#)

BibSonomy is a system for sharing bookmarks and lists of literature. When discovering a bookmark or a publication on the web, you can store it on our server. You can add tags to your entry to retrieve it more easily. This is very similar to the bookmarks/favorites that you store within your browser. The advantage of BibSonomy is that you can access your data from wherever you are. Furthermore, you can discover more bookmarks and publications from your friends and other people.

This page shows you the latest updates of BibSonomy. Why don't you just try it yourself? After a free [registration](#), you can organise your own bookmarks and publications, and discover related entries.

busy tags

[dblp](#)
[lang:de](#)
[programming](#)
[DRM](#)
[boomerang](#)
[reference](#)
[java](#)
[wood](#)
[shop](#)
[ir](#)

bookmarks

[previous](#) | [next](#)

NICHTLUSTIG.DE

to humor cartoon lang:de as public by schmitz and 1 other person on 2005-12-18 23:13:12.0

BibTeX

[previous](#) | [next](#)

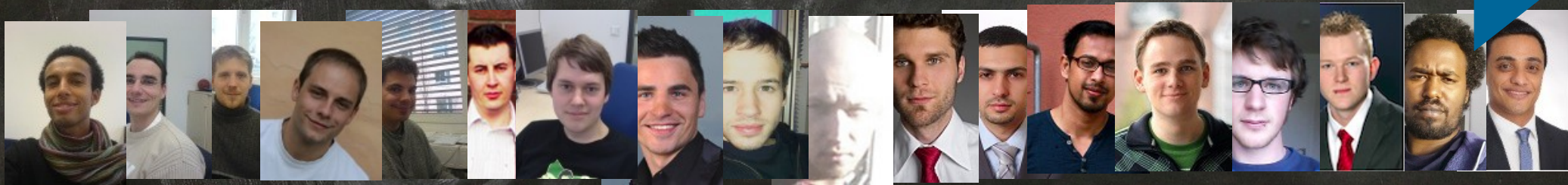
Comparison of visualizations in formal concept analysis and correspondence analysis

K. E. Wolff and S. Gabler. *Visualization of categorical data*. (to appear)

History



2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016





ENDOWED CHAIR OF THE HERTIE FOUNDATION

Knowledge and Data Engineering

ELECTRICAL ENGINEERING & COMPUTER SCIENCE, UNIVERSITY OF KASSEL



Information
School.



History



PRALINÉS - VARIÉTÉ

Feine Spezialitäten aus feinsten Chocolates



user ▾

@jaeschke

tag(s)



tag

user

group

author

concept

BibTeX key

search

search:jaeschke

in group

- 20dc13

- bibsonomy_dev

- iccs

- kde

- l3s

- nepomuk

- stair

@jaeschke

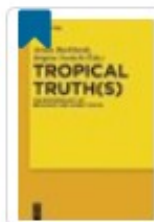


publications

1098



Tropical Truth(s): The Ep



A. Burkhardt, and

5 days ago by

epistem

★★★★★ (0)

Why does deep and che



H. Lin, and M. Te

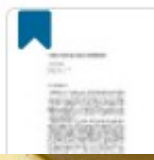
arxiv:1608.08225

6 days ago by

deep

★★★★★ (0)

Web Archiving: Issues a



J. Masanès. pag

12 days ago by

actions for all displayed publications:

edit own posts

filter

display: all, [publications only](#)

publications per page: [5](#) [10](#) [20](#)

[50](#) [100](#)

export

[RSS](#), [BibTeX](#), [RDF](#), [more...](#)

sort

criterion: ([date](#) | [title](#))

order: ([ascending](#) | [descending](#))

[help for advanced sorting...](#)

clipboard

[pick all](#), [unpick all](#)

private documents

upload document

Durchsuchen...

Keine Datei ausgewählt.

The uploaded documents are stored temporarily. After clicking save, they are added to the post. The documents will be available only to you and members of your groups.

added files

- Supplementary Materials.pdf

Comments and Reviews

rating distribution



average user rating



5.0 out of 5.0 based on 1 review



anonymous 5 years ago



good summary of research and technical details around the bibsonomy tool

Users



edit your tags and concepts

rename/replace tags

On every post that contains all the tags from the first box these tags will be replaced by those from the second box.

replace these tag(s)

by these tags

☐ replace tags in concepts (replacing tags in concepts only works, when exactly one tag is being replaced by one other tag)

replace

reset

add subtags to concepts

supertag

subtag

insert

reset

delete subtags from concepts

supertag

group / 20dc13 / curriculum vitae

Curriculum Vitae

Grouppage of the group 15th Discovery Challenge



15th Discovery Challenge

Members



@folke



@hotho



@jaeschke



@juergen...



@sdo



You shall not pass: detecting malicious users at registration time

Full Text:  PDF

Authors: [Christian Kater](#) [Leibniz University Hannover](#)
[Robert Jäschke](#) [Leibniz University Hannover](#)

Published in:


· Proceeding
[OnSt '16](#) Proceedings of the 1st International Workshop on Online Safety,
Trust and Fraud Prevention

Article No. 2

ACM New York, NY, USA ©2016

[table of contents](#) ISBN: 978-1-4503-4364-0 doi>[10.1145/2915368.2915370](#)



 2016 Article



[Bibliometrics](#)

- Downloads (6 Weeks): 8
- Downloads (12 Months): 30
- Downloads (cumulative): 30
- Citation Count: 0

ACM DL DIGITAL LIBRARY

You shall not

Full Text: 

Authors: Chr
Rok

Published in:

Proceeding

OnSt '16 Proc

Trust and Fra

Article No. 2

ACM New York, N

[table of contents](#) I

save publication

BibSonomy :: home

open myBibSonomy

save bookmark

You shall no

Full Text: 

Authors: Chr
Rok

Published in:

· Proceeding

OnSt '16 Proc

Trust and Fra

Article No. 2

ACM New York, N

[table of contents](#) I

save publication

edit your publication post

general information

entry type

inproceedings

suggestions are available

kater2016shall

title

You Shall Not Pass: Detecting Malicious Users at Registration Time

author(s)

Kater, Christian
Jäschke, Robert

☒ I am an author

editor(s)

year

2016

tags - describe the post

tags

2016 bibsonomy fraud myown spam

space separated

recommendation

web

social

bibsonomy

blog

bibsonomynews



You shall not pass

Full Text: Authors: Chr
Rok

Published in:

Proceeding

OnSt '16 Proc

Trust and Fra

Article No. 2

ACM New York, N

[table of contents](#) I

save publication

BibSonomy edit your publication post

general information

home myBibSonomy add post groups popular genealogy admin

myBibSonomy / You Shall Not Pass: De...



You Shall Not Pass: Detecting Malicious Users at Registration Time

C. Kater, and R. Jäschke. *Proceedings of the 1st International Workshop on Online Safety, Trust and Fraud Prevention, page 2:1--2:6. New York, NY, USA, ACM, (June 2016)*

Abstract

Spam is a widespread problem for many online services. The use case in this paper is the social bookmarking system BibSonomy, which received over 150 times more registrations from spam users than from normal users over the last ten years. A common approach to fight spam is to use machine learning to classify the users into good or malicious users. ... [\(more\)](#)

Links and resources

DOI: 10.1145/2915368.2915370

URL: <http://doi.acm.org/10.1145/2915368.2915370>

additional links:

+ add URL

OpenURL: <https://getinfo.de/app/openurl/search?date=2016&id=doi%3A10.1145%...>

BibTeX key: kater2016shall

search on:

tags - de

Comments and Reviews ★★★★★ (0)

There is no review or comment yet. [You can write one!](#)

Tags

fraud myown 2016 spam bibsonomy sys.relevantfor:is

You Shall Not Pass: Detecting Malicious Users at Registration Time

Christian Kater
University of Bremen
ckater@informatik.uni-bremen.deRobert Jäschke
University of Bremen
jaeschke@informatik.uni-bremen.de

ABSTRACT

Spam is a widespread problem for many online services. The use case in this paper is the social bookmarking system BibSonomy, which received over 150 times more registrations from spam users than from normal users over the last ten years. A common approach to fight spam is to use machine learning to classify the users into good or malicious users. ... [\(more\)](#)

KEYWORDS

spam, machine learning, social bookmarking

1. INTRODUCTION

Spam is a widespread problem for many online services. The use case in this paper is the social bookmarking system BibSonomy, which received over 150 times more registrations from spam users than from normal users over the last ten years. A common approach to fight spam is to use machine learning to classify the users into good or malicious users. ... [\(more\)](#)

Spam is a widespread problem for many online services. The use case in this paper is the social bookmarking system BibSonomy, which received over 150 times more registrations from spam users than from normal users over the last ten years. A common approach to fight spam is to use machine learning to classify the users into good or malicious users. ... [\(more\)](#)

kater2016detecting.p

add a document

You shall not pass

Full Text: 

Authors: Chr

Rok

Published in:

· Proceeding

OnSt '16 Proc

Trust and Fra

Article No. 2

ACM New York, N

[table of contents](#) I

save publication

BibSonor edit your publication post

general information

[home](#) [myBibSonomy](#) [add post](#) [groups](#) [popular](#) [genealogy](#) [admin](#)[home](#) / [myBibSonomy](#) / You Shall Not Pass: De...

You Sh

C. Kater, a

page 2:1--2

Robert Jäschke

[Home](#)[Publications](#)[Events](#)[Projects](#)[Teaching](#)[Code](#)[Data](#)[Links](#)

Abstract

Spam is a widespread problem in the social bookmarking services. A common approach to fight good or malicious users.

Links and resource

DOI

URL

additional link:

OpenURL

BibTeX key

search on

Comments and Re

reci There is no review or co

Tags

[fraud](#)[myown](#)[2016](#)[spam](#)[bibsonomy](#)[sys.relevantfor:is](#)

Publications

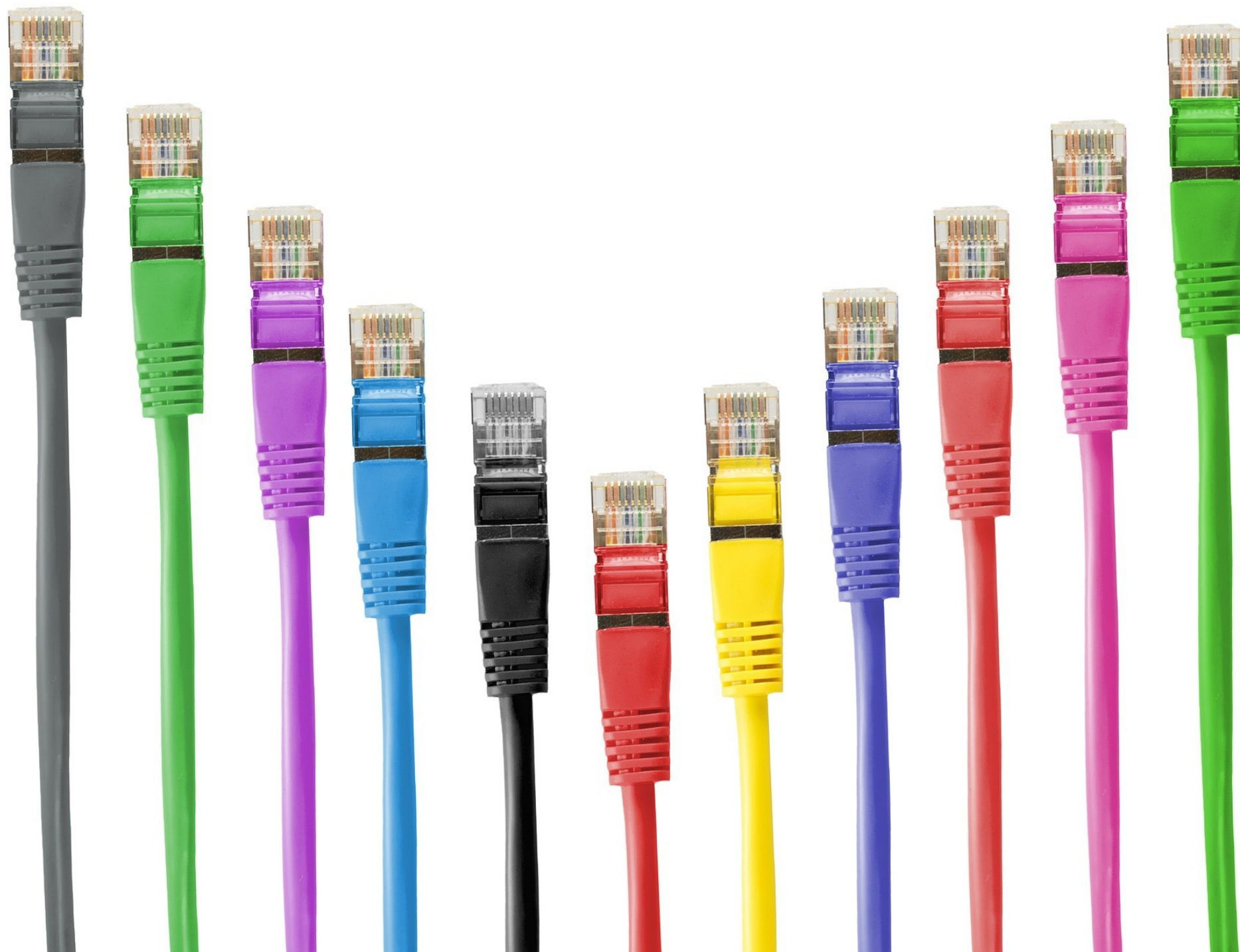
My publications are also available in [BibSonomy](#) (as BibTeX) and [DBLP](#). This list is generated with [Jekyll](#) and [bibsonomy-ruby](#). ([color legend](#))

2016

- Bergmann, J., Hadgu, A.T., Jäschke, R.: **Tweeting in times of exposure: A mixed-methods approach for exploring patterns of communication related to business scandals on Twitter**. Proceedings of the Workshop on Natural Language Processing and Computational Social Science. , Hannover, Germany (2016). [[PDF](#) | [BibTeX](#) | [BibSonomy](#)]
- Hadgu, A.T., Lotze, N., Jäschke, R.: **Telling English Tweets Apart: the Case of US, GB, AU**. Proceedings of the Workshop on Natural Language Processing and Computational Social Science. , Hannover, Germany (2016). [[PDF](#) | [BibTeX](#) | [BibSonomy](#)]
- Kater, C., Jäschke, R.: **You Shall Not Pass: Detecting Malicious Users at Registration Time**. Proceedings of the 1st International Workshop on Online Safety, Trust and Fraud Prevention. pp. 2:1–2:6. ACM, Hannover, Germany (2016). [[PDF](#) | DOI:10.1145/2915368.2915370 | [URL](#) | [BibTeX](#) | [BibSonomy](#)]




- thinking about hidden features
- removing unused features is sometimes necessary
- there is never a shortage of ideas for good features
- explaining (new) features is important
- having an HCI expert is important




Integration with other Websites and Services

BibSonomy integrates well with other websites, content management systems, and reference managers. Furthermore, you can connect BibSonomy to your own software using one of our API (application programming interface) clients. The following list describes ways to integrate BibSonomy with various systems and technologies - ranging from code snippets to complete and powerful, easy-to-install plugins or API clients.

- Integration with other Websites and Services
 - Browser
 - Chrome / Chromium
 - Firefox
 - Add on
 - Search
 - Zotero
 - Safari
 - Programs
 - Citavi
 - Emacs
 - JabRef
 - BibSonomy Plugin
 - Extract JabRef layouts from Jar files
 - KBibTeX
 - Sublime Text
 - TeXlipse
 - Websites
 - Confluence
 - GoogleDocs
 - GoogleScholar
 - Moodle
 - Your own website
 - Bookmark Links
 - Tag Clouds
 - Post Lists
 - JSON feed
 - TYPO3
 - Wordpress
 - XWiki
 - Zope
 - Libraries
 - Digital Libraries
 - Discovery Service
 - OpenURL resolver
 - Applications
 - Bookmark Exporter
 - Link checker
 - RSS
 - REST API
 - Java
 - PHP
 - Python
 - Others
 - Android
 - OAuth
 - Search Engine

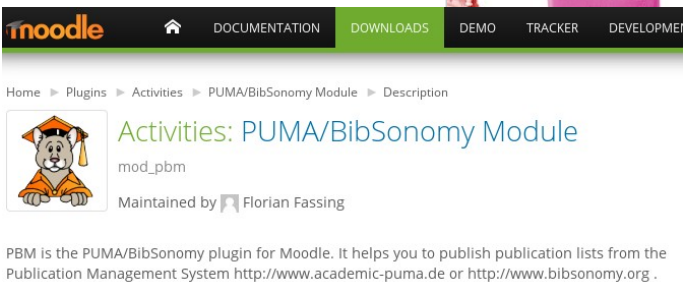


Add-ons
EXTENSIONS THEMES COLLECTIONS MORE...

 **BibSonomy Buttons** 1.6.2
by [BibSonomy](#)



Three useful BibSonomy buttons: myBibSonomy as well as post bookmark and post publication.

[+ Add to Firefox](#)



moodle [DOCUMENTATION](#) [DOWNLOADS](#) [DEMO](#) [TRACKER](#) [DEVELOPMENT](#)

Home ► Plugins ► Activities ► PUMA/BibSonomy Module ► Description

 **Activities: PUMA/BibSonomy Module**
mod_pbm
Maintained by  Florian Fassing

PBM is the PUMA/BibSonomy plugin for Moodle. It helps you to publish publication lists from the Publication Management System <http://www.academic-puma.de> or <http://www.bibsonomy.org>.



TYPO3

[Home](#) [About](#) [TYPO3 CMS](#) [Documentation](#) [Extensions](#) [Support](#)

[typo3.org](#) > [Extensions](#) > Extension Repository

 **PUMA/BibSonomy CSL**

Add a list of publications and tag clouds from BibSonomy or PUMA to your TYPO3 pages. Style your publication lists using CSL stylesheets (<http://citationstyles.org>).

All sites supported by our scrapers

The following list contains all supported catalogs whose publications can be extracted for you with the "[postPublication](#)"-Button or the [Browser extensions](#). The links make a good starting point for importing publications of interests into BibSonomy.

Bibliothek Kassel | IEEEExplore Journals | IEEEExplore Books | IEEEExplore Standards | SpringerLink | ScienceDirect | PubMed | PubMedCentral | SLAC National Accelerator Laboratory | ACM Digital Library | AIFB | University Library (UB) Karlsruhe | arXiv | Ingentaconnect | Librarything | The SAO/NASA Astrophysics Data System | AIP Scitation | JOURNAL OF THE AMERICAN SOCIETY OF HEMATOLOGY | MathSciNet | InterScience | IOP | PROLA | BibSonomy | IEEE Computer Society | Amazon | PLoS | Nature | Blackwell Synergy | University of Trier Digital Bibliography & Library Project | BioMed Central | Worldcat | Springer | <http://www.acs.org/> | AnthroSource | BMJ | Ed/ITLib | Informaworld | Cambridge Journals | Liebert Online | National Bureau of Economic Research | USENIX | International Union of Crystallography | Optical Society of America | Psy CONTENT | Royal Society Publishing | PNAS | Science Magazine | Jstor | Education Resources Information Center | IWA Publishing | Journal of Machine Learning Research | Association for Computational Linguistics | J-Stage | Annual Reviews | Project MUSE | SSRN | SCOPUS | Meta Press | CiteSeerX | Open Repository | Pion | CiteUlike | American Meteorological Society | LIIN | Wormbase | Google Scholar | Google Patente | SciELO Scientific Electronic Library Online | D-Lib | Cell | Cases Journal | Elsevier: Medical publishers, online journals | Astronomy and Astrophysics | Journalogy (Microsoft Academic Search) | INSPIRE | Taylor & Francis Online | Google Books | Journal of Applied Physiology | American Thoracic Society Journals | The New England Journal of Medicine | Sage Journal | The Journal of Clinical Investigation | Stanford InfoLab Publication Server | Royal Society of Chemistry | The FASEB Journal | The Journal of Neuroscience | Development | The Open University | Association for the Advancement of Artificial Intelligence | American Psychological Association | BMJ Open | Pharmacognosy Research | EconStor | JCB | American Psychological Society | Mendeley | Hindawi Publishing Corporation | American Society for Microbiology | First Monday | IGI Global | De Gruyter | Morgan & Claypool Publisher | MDPI - Open Access Publishing | Oxford Journals | THE LANCET | World Scientific | Now The essence of Knowledge | Genome Biology | American Journal of PUBLIC HEALTH | Cold Spring Harbor Perspectives in Biology | RWTH Aachen University | U.S. Department of energy - Office of Scientific and technical information | scientific | Copac National, Academic and Specialist Library Catalogue | cinii Articles | Aha Journals | Karger medical and scientific publishers | Cold Spring Harbor Perspectives in Biology | SPIE Digital Library | Bio One Research Evolved | The University of Chicago Press Journals | Refdoc | Pediatrics official journal of the american academy of pediatrics | Journal of Neurological Surgery Part A | The Journal of American Medical Association | Astronomy and Astrophysics | Library of Congress | Elsevier | Compare billion project of the Russian Federation Ministry of Culture and Public Initiative | ContentNegotiationDOIScraper | EprintsScraper | UnAPIScraper | Highwire Scraper Collection | CoinsScraper | DublinCoreScraper

The details of each scraper will be explained in the following. For more information, have a look at the [help pages](#).



format

- please select -

posts

- please select -

BibTeX

CSL

EndNote

MSOffice XML

APA: HTML according American Psychological Association (APA) guidelines

APA: RTF according American Psychological Association (APA) guidelines

BibO RDF

BibTeXML

Boxed HTML

BuRST

CV BibSonomy

Chicago Manual of Style

DBLP XML

DIN 1505

DIN 1505 MMS

DIN 1505 MMS, ENDNOTE, NOTE, YEAR

DIN 1505 MMS, ENDNOTE, YEAR

DIN 1505 MMS, YEAR

DIN 1505, YEAR

By selecting one of the formats you

Favorite layouts



.bib

BibTeX

BibTeX is a popular k
bibliography{FILENA



.html

CSL

The Citation Style La
compatible with most



.txt

EndNote *

This is an output in E

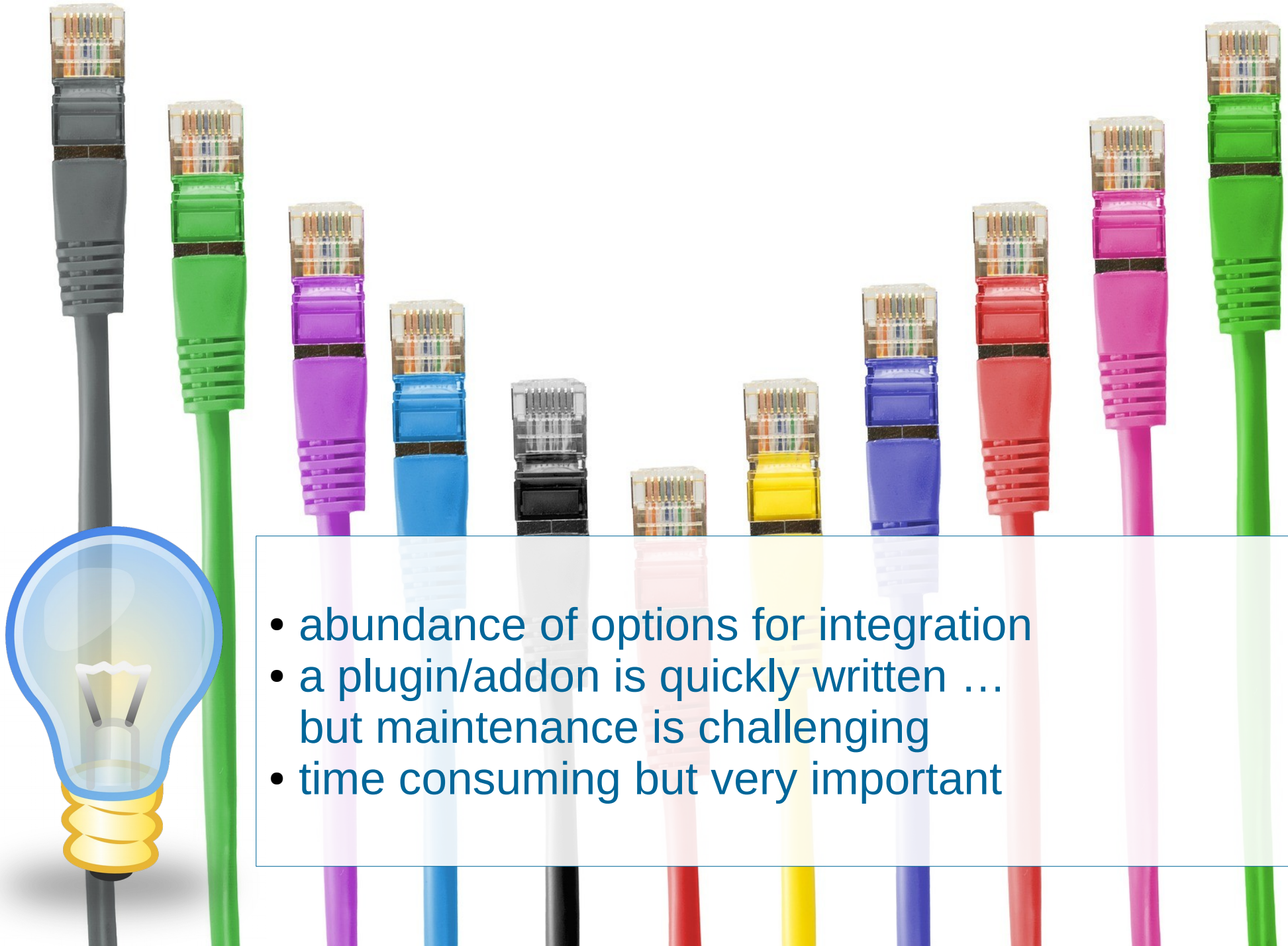


.xml

MSOffice XML *

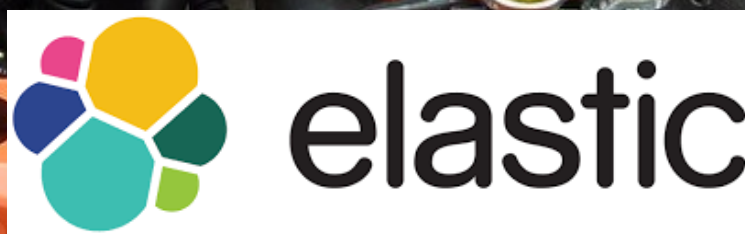
A format to import into

Microsoft Bibliography.



- abundance of options for integration
- a plugin/addon is quickly written ... but maintenance is challenging
- time consuming but very important







- invest time in proper architecture early on
- migration is challenging
- many invisible improvements
- do not fear change

BibSonomy::home

search [home](#) :: [tags](#) :: [popular](#)username: password: [help](#)[login](#) :: [register](#)

BibSonomy is a system for sharing bookmarks and lists of literature. When discovering a bookmark or a publication on the web, you can store it on our server. You can add tags to your entry to retrieve it more easily. This is very similar to the bookmarks/favorites that you store within your browser. The advantage of BibSonomy is that you can access your data from wherever you are. Furthermore, you can discover more bookmarks and publications from your friends and other people.

This page shows you the latest updates of BibSonomy. Why don't you just try it yourself? After a free [registration](#), you can organise your own bookmarks and publications, and discover related entries.

busy tags

[dblp](#)[lang:de](#)[programming](#)[DRM](#)[boomerang](#)[reference](#)[java](#)[wood](#)[shop](#)[ir](#)

bookmarks

[previous](#) | [next](#)

NICHTLUSTIG.DE

to humor cartoon lang:de as public by schmitz and 1 other person on 2005-12-18 23:13:12.0

-- . - . - .

BibTeX

[previous](#) | [next](#)**Comparison of visualizations in formal concept analysis and correspondence analysis**

K. E. Wolff and S. Gabler. *Visualization of categorical data*, (to appear)



bookmarks (410) [RSS](#) [XML](#)

<< < 1 | 2 | 3 > >>

Generic SOAP Client

This generic SOAP client allows you to access web services using a web browser. It performs dynamic bindings and executes methods at remote web service...

to [test wsdl online soap client](#) by [jaeschke](#) on Jul 1, 2008, 12:37 PM

Popsongs und ihre Hintergründe - Songarchiv

to [rock history background geschichte song music pop hintergrund](#) by [jaeschke](#) and [1 other person](#) on Jun 28, 2008, 12:43 AM

Popsongs und ihre Hintergründe - Songarchiv

to [song hintergrund pop geschichte history music background rock](#) by [jaeschke](#) and [1 other person](#) on Jun 28, 2008, 12:37 AM

Wordle - Beautiful Word Clouds

Wordle is a toy for generating "word clouds" from text that you provide. The clouds give greater prominence to words that appear more frequently in the sou...

to [cloud word wordle](#) by [jaeschke](#) and [10 other people](#) on Jun 23, 2008, 9:04 PM

ONJava.com -- Advanced Configuration of the Spring MVC Framework

In this article I will present some configuration tips for the Spring MVC framework that could help you manage multiple instances of your Spring-based web ...

to [web configuration programming mvc spring](#) by [jaeschke](#) on Jun 19, 2008, 3:36 PM

DadaDada

publications (208) [RSS](#) [BibTeX](#) [RDF](#) [more](#)

<< < 1 | 2 | 3 > >>

Detecting Communities via Simultaneous Clustering of Graphs and Folksonomies

Akshay [Java](#) and Anupam [Joshi](#) and Tim [Finin](#) *WebKDD 2008 Workshop on Web Mining and Web Usage Analysis, August 2008. To Appear.*
to [detection clustering community](#) by [jaeschke](#) and [1 other person](#) on Jul 12, 2008, 1:03 PM
[BibTeX](#)

TRIAS - An Algorithm for Mining Iceberg Tri-Lattices

Robert [Jäschke](#) and Andreas [Hotho](#) and Christoph [Schmitz](#) and Bernhard [Ganter](#) and Gerd [Stumme](#) *ICDM '06: Proceedings of the Sixth International Conference on Data Mining, page907--911. Washington, DC, USA, IEEE Computer Society, (2006)*
to [trias myown 2006](#) by [jaeschke](#) and [10 other people](#) on Jun 23, 2008, 12:23 PM
[URL](#) | [BibTeX](#)

Logsonomy - Social Information Retrieval with Logdata

Beate [Krause](#) and Robert [Jäschke](#) and Andreas [Hotho](#) and Gerd [Stumme](#) *HT '08: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia, page157--166. New York, NY, USA, ACM, (2008)*
to [engine information l3s logsonomy for:nepomuk wp5 social retrieval analysis network search](#) by [jaeschke](#) and [2 other people](#) on Jun 20, 2008, 2:42 PM
[URL](#) | [BibTeX](#)

Logsonomy — A Search Engine Folksonomy

Robert [Jäschke](#) and Beate [Krause](#) and Andreas [Hotho](#) and Gerd

filter:

- relations

conference ← [ecai](#) [gvd](#) [iccs](#) [icdm](#) [icfca](#)
geo ← [gps](#) [map](#) [utm](#)
howto ← [manual](#) [reference](#) [tutorial](#)
location ← [anhalt](#) [berlin](#) [bitterfeld](#) [celle](#) [dresden](#)
[europe](#) [frankfurt](#) [hannover](#) [kassel](#) [london](#)
[magdeburg](#) [ort](#) [sachsen](#) [sachsen_anhalt](#)
[tübingen](#) [wittenberg](#)
ort ← [location](#)
programming ← [ada](#) [c](#) [fortran](#) [java](#) [lisp](#) [perl](#) [ruby](#)
protocol ← [ftp](#) [http](#) [smtp](#)
researcher ← [devadze](#) [shannon](#) [turing](#)
science ← [chemistry](#) [math](#)
software ← [apache](#) [beagle](#) [cocoa](#) [debian](#)
[eclipse](#) [firefox](#) [haystack](#) [nextstep](#)
[photoshop](#) [protege](#) [thunderbird](#) [weka](#)
[wine](#) [word](#) [x11](#) [zope](#)
sport ← [basketball](#) [football](#) [volleyball](#)

- tags

([alpha](#) | [freq](#)) ([cloud](#) | [list](#)) ([minfreq](#) [1](#) | [2](#) | [5](#))
[12](#) [2005](#) **[2006](#)** [2007](#) [2008](#) [129a](#) [3-mode](#) [3d](#)
[3elements](#) [3mode](#) [abmahnen](#) [absolute](#) [abuse](#) [acl](#)
[acronym](#) [actor](#) [ada](#) [address](#) [admin](#) [adsense](#)
[advertisement](#) [aggregation](#) [agrawal](#) [ai](#) [air](#) [ajax](#)
[algebra](#) **[algorithm](#)** [all](#) [alt](#) [altitude](#) [amateurfunk](#)
[amatex](#) [amazon](#) [amine](#) **[analysis](#)** [analyze](#)
[animal](#) [animation](#) [anleitung](#) [annotation](#) [anti](#)
[antiquariat](#) [antiquarisch](#) [anwalt](#) [apache](#) [api](#) [apple](#)
[architecture](#) [architektur](#) [archive](#) [ard](#) [art](#) [article](#)
[association](#) [atom](#) [audio](#) [auskunft](#) [auto](#) [aviation](#)
[award](#) [awk](#) [axis](#) [axis2](#) [background](#) [bahn](#) [balloon](#)
[bank](#) [basteln](#) [batch](#) [bayes](#) [beagle](#) [berlin](#)
[bibliographic](#) [bibsonomy](#) [bibtex](#) [bienen](#) **[binary](#)**
[bing](#) [bin](#) [bing](#) [binh](#) [bin](#) [black&white](#) [blind](#)

BibSonomy :: ▾ user :: jaeschke ::

A blue social bookmark and publication sharing system.

[Home](#) ▾ [myBibSonomy](#) [post bookmark](#) [post publication](#) [tags](#) [authors](#) [relations](#) ▾ [groups](#) [popular](#)

<<

bookmarks (697)[RSS](#) [XML](#)

<< < 1 | 2 | 3 > >>

[edit](#)**LaTeX Beamer Theme Previews « Jinto Reedwine's Blog**

I have been using LaTeX beamer for a little over a year now to do my "PowerPoint" like presentations for math and computer science classes. I really like b...

to [beamer latex preview theme](#) by [jaeschke](#) on Jun 28, 2010, 8:33 PM

[edit](#) | [delete](#) | [spam](#)**PUMA - erste stabile Version jetzt online verfügbar**

Das im August gestartete DFG-Projekt PUMA - Akademisches Publikationsmanagement ist mittlerweile in einer ersten Version online verfügbar.

to [bibliothek blog kassel projekt puma](#) by [jaeschke](#) on Jun 28, 2010, 2:49 PM

[edit](#) | [delete](#) | [spam](#)**Feature of the week: New help pages**

to [bibsonomy bibsonomynews blog fotw help](#) by [jaeschke](#) on Jun 28, 2010, 2:44 PM

[edit](#) | [delete](#) | [spam](#)**Coranac » Off on a tangent : a look at****publications** (392)[RSS](#) [BibTeX](#) [RDF](#) [more](#)

<< < 1 | 2 | 3 > >>

[edit](#) | [pick](#) | [unpick](#)**Query Logs as Folksonomies**

Dominik [Benz](#) and Andreas [Hotho](#) and Robert [Jäschke](#) and Beate [Krause](#) and Gerd [Stumme](#) *Datenbank-Spektrum*10(1):15–24jun 2010.

to 2010 folksonomy log logsonomy myown query by [jaeschke](#) and 1 other user on Jun 14, 2010, 10:11 AM

[pick](#) | [edit](#) | [delete](#) | [gold-standard](#) | [URL](#) | [BibTeX](#) | [openURL](#) | [spam](#)**Learning optimal ranking with tensor factorization for tag recommendation**

Steffen [Rendle](#) and Leandro Balby [Marinho](#) and Alexandros [Nanopoulos](#) and Lars [Schmidt-Thieme](#) *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, page 727–736. New York, NY, USA, ACM, (2009)*

to [folksonomy ranking recommender tag tagging tensor](#) by [jaeschke](#) and 5 other users on Jun 14, 2010, 10:09 AM

[pick](#) | [edit](#) | [delete](#) | [gold-standard](#) | [URL](#) | [DOI](#) | [BibTeX](#) | [openURL](#) | [spam](#)**A fast string searching algorithm**

Robert S. [Boyer](#) and J. [Strother Moore](#) *Communications of the ACM*20(10):762–772oct 1977.

to [algorithm boyer match moore search string](#) by [jaeschke](#) and 2

filter: [+ relations](#)

(show all | hide all)

[+ similar users more...](#)[- tags](#)

(alpha | freq) (cloud | list) (minfreq 1 | 2 | 5)

2006 2008 algorithm analysis
annotation bibsonomy
bibsonomynews bibtex binary blog
book bookmarking buch bunker
classification clustering collaborative
community concept dagsocial data
database desktop detection diploma
engine fca **folksonomy** for:kde
formal fotw fun geo google gps
graph history howto
iccs_example information java
kassel knowledge l3s latex learning
lehre linux location manual map
math memory mining music myown
mysql **network** ontology ort
pagerank photo privacy



1

👍👍👍👍👍 (0)

1

★★★★★ (0)

1

☆☆☆☆☆ (0)

1

★★★★★ (0)

1

☆☆☆☆☆ (0)

2

★★★★☆ (1)

1



0

★★★★★ (0)

•

★★★★★ (0)

★★★★★ (0)

★★★★★ (0)

★★★★★ (0)

1

(show all | hide all)

↓ to* ← tobuy todo toread towatch

```
hotho stumme dbenz schmitz macedet flint63
dolefulrabbit sdo beate brightbyte
```

bookmarks 1228



OpenCitations - Home

1



<http://opencitations.net/>

🕒 2 days ago by @jaeschke

🔖 citation data dataset open

★★★★★ (0)



publications 1082



Tropical Truth(s): The Epistemology of Metaphor and ot...

1



A. Burkhardt, and B. Nerlich (Eds.). *De Gruyter, (2010)*

🕒 5 days ago by @jaeschke

🔖 dh epistemology language metaphor text trope

★★★★★ (0)



Handbuch Open Science – Wikibooks, Sammlung freie...

1



https://de.wikibooks.org/wiki/Handbuch_Open_Science

🕒 23 days ago by @jaeschke

🔖 handbuch open science wikibooks

★★★★★ (0)



Why does deep and cheap learning work so well?

3



H. Lin, and M. Tegmark. (2016) *cite arxiv:1608.08225* Comment: 14 pages, 3 figs.

🕒 6 days ago by @jaeschke

🔖 deep learning machine network neural

★★★★★ (0)



Comparison of research networking tools and research...

1



https://en.wikipedia.org/wiki/Comparison_of_research_...

🕒 2 months ago by @jaeschke

🔖 gaw network profile researcher

★★★★★ (0)



Web Archiving: Issues and Methods

1



J. Masanès. *page 1--53. Springer, Berlin/Heidelberg, (2006)*

🕒 12 days ago by @jaeschke

🔖 archive web

★★★★★ (0)



Hermitage: Testing the "I" in ACID — Martin Kleppman...

1



<https://martin.kleppmann.com/2014/11/25/hermitage-tes...>

🕒 2 months ago by @jaeschke

🔖 acid base database isolation nosql sql

★★★★★ (0)



Determining the Optimal Session Interval for Transactio...

2



S. Wakeling, and P. Clough. *page 703--708. Springer, (2016)*

🕒 14 days ago by @jaeschke

🔖 catalogue digital dl interval library log session

★★★★★ (0)



@jaeschke

Robert Jäschke



You must be signed in to add this user to your friend list.

sign in

discussion

show posts recently discussed by jaeschke

concepts

similar users

(more...)

@macdet @schmitz @flint63 @dolefulrabbit
@stumme @hotho @beate @dbenz @sdo
@brightbyte

shared groups

tags

alexandria analysis archive

bibsonomy bibsonomynews

blog bookmarking citation collaborative

collective community computing concept data

dataset fca folksonomy fotw geo

graph information intelligence java library

map mining myown network ontology

open paper programming publication

recommender research science

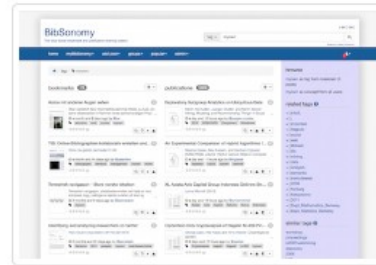
scientometrics search semantic sna

social survey tag tagging toread twitter

visualization web wikipedia



BibSonomy denotes a **social bookmarking system** as well as the underlying **technical platform**. The system enables users to share and annotate bookmarks and publications. This page is the entry point for developers interested in the software platform underlying **BibSonomy** and **PUMA**.



The web application is based on Java, backed-up by a MySQL database and Elasticsearch for fulltext search. The BibSonomy platform is developed by the **KDE group** of the University of Kassel, the **DMIR group** of the University of Würzburg, the **L3S Research Center**, and **HothoData**.

Development

- Overview
- API Documentation
- Maven Repository
- REST Clients for Java, PHP, Python and Ruby
- Contributing Scrapers

Integration

- PUMA
- TYPO3 Extension
- WordPress Plugin
- TeXlipse Plugin
- JabRef Plugin
- List of Services

Research

- Citing BibSonomy
- Data Dumps
- Our Publications
- "BibSonomy dataset"
- Discovery Challenges 2008 and 2009

What is BibSonomy?

- Getting Started
- Browser Buttons
- Help
- Feature of the Week
- Scientific Perspective

About & Contact

- Team
- Terms of Use
- Contact
- Reporting Issues
- Donating

News

- [Blog](#)
- [@BibSonomyCrew](#)
- [#bibsonomy](#)
- [Google+ Community](#)
- [Mailing List](#)

<http://dev.bibsonomy.org/>



BibSonomy Development Portal

The features for the next release (3.7.0) are documented in the [Release Log](#).

Starting as a Developer

Start with reading the description on [how to set up a BibSonomy workspace](#). If you are experiencing any problems, contact one of the [main developers](#). The source code is regularly released on [bibsonomy.org](#) - what this means for you is documented in the [release process](#).

Tutorials

- Best Practices
- How to use the version control system
- Errors and Fixed
- Eclipse Tips and Tricks
- User Support - how to help users that have problems with BibSonomy.
- SSL-Certificates

Module Documentation

This section provides information about some of the modules of BibSonomy.

- general Configuration
- modified programming libraries

Model

- How do the [publication hashes](#) work?

Database

- changing the [Database Schema](#)
- assignment of the [Chains of Responsibility](#)

Webapp

- Internationalization
- EL, JSP & JSTL Documentation
- Infos regarding jQuery updates
- Spring Security
- UriRewriteFilter
- Structured Data Markup

Other

- Clicklog
- Help
- Lucene
- Shared Resource Search
- Pingback: Pingback and Trackback
- Scraper
- Sync
- Person Names
- Changed Posts
- OpenSocial
- SystemTags
- SystemTagsArchitecture

Planned Developments and Discussions

Roadmap

- History
- Accessing Documents
- Community Posts
- Groups
- UpdateScript
- Development Discussion





- plan for times with limited resources
- invest time in community building
- do not underestimate complexity of development



The background of the slide is a photograph of a NASA Mission Control Center. It features a large curved wall of monitors displaying various data, including a view of Earth from space, orbital paths, and technical readouts. The room is filled with computer equipment and has NASA logos and mission patches on the walls.

load balancing

replication

backup

geographical
distribution

protection

monitoring

notification

maintenance

user relations

communication

```
bibsonomy@gandalf:~$ uptime
10:03:44 up 1041 days, 21:05,  4 users,  load average: 0.00, 0.01, 0.05
bibsonomy@gandalf:~$ date
Do 8. Sep 10:04:19 CEST 2016
```




New registrations

Admin: Spammer

Admin: Unsure

Admin: No Spammer

Classifier: Spammer

Classifier: Spammer (U)

Classifier: No Spammer (U)

Classifier: No Spammer

Modified BibTeX Users

(show 100 - 500 - 1000 entries)

Spammer?	Username	Name	IP	E-Mail	Regdate	Algorithm	Mode	Conf.	Updated
no ? yes					Sep 28, 2016 10:54 PM	Logistic_gm	D	0.574	Sep 28, 2016 10:59 PM
no ? yes					Aug 18, 2016 7:16 AM	Logistic_gm	D	0.569	Sep 28, 2016 10:32 PM
no ? yes					Sep 28, 2016 9:36 PM	Logistic_gm	D	0.658	Sep 28, 2016 9:41 PM
no ? yes					Aug 10, 2016 2:45 PM	Logistic_gm	D	0.671	Sep 28, 2016 9:21 PM
no ? yes					Sep 28, 2016 7:37 PM	Logistic_gm	D	0.636	Sep 28, 2016 8:19 PM
no ? yes					Apr 21, 2016 11:03 AM	Logistic_gm	D	0.667	Sep 28, 2016 8:03 PM

Dear Webmaster,

I am the webmaster of [www. \[redacted\] .au](http://www. [redacted] .au). It would be greatly appreciated if you remove links to my website. My Links can be found on these following pages

<http://www.bibsonomy.org/user/lowcostseo11?lang=ru>
<http://www.bibsonomy.org/user/lowcostseo11?resourcetype=publication>
<http://www.bibsonomy.org/user/lowcostseo11?resourcetype=bookmark>

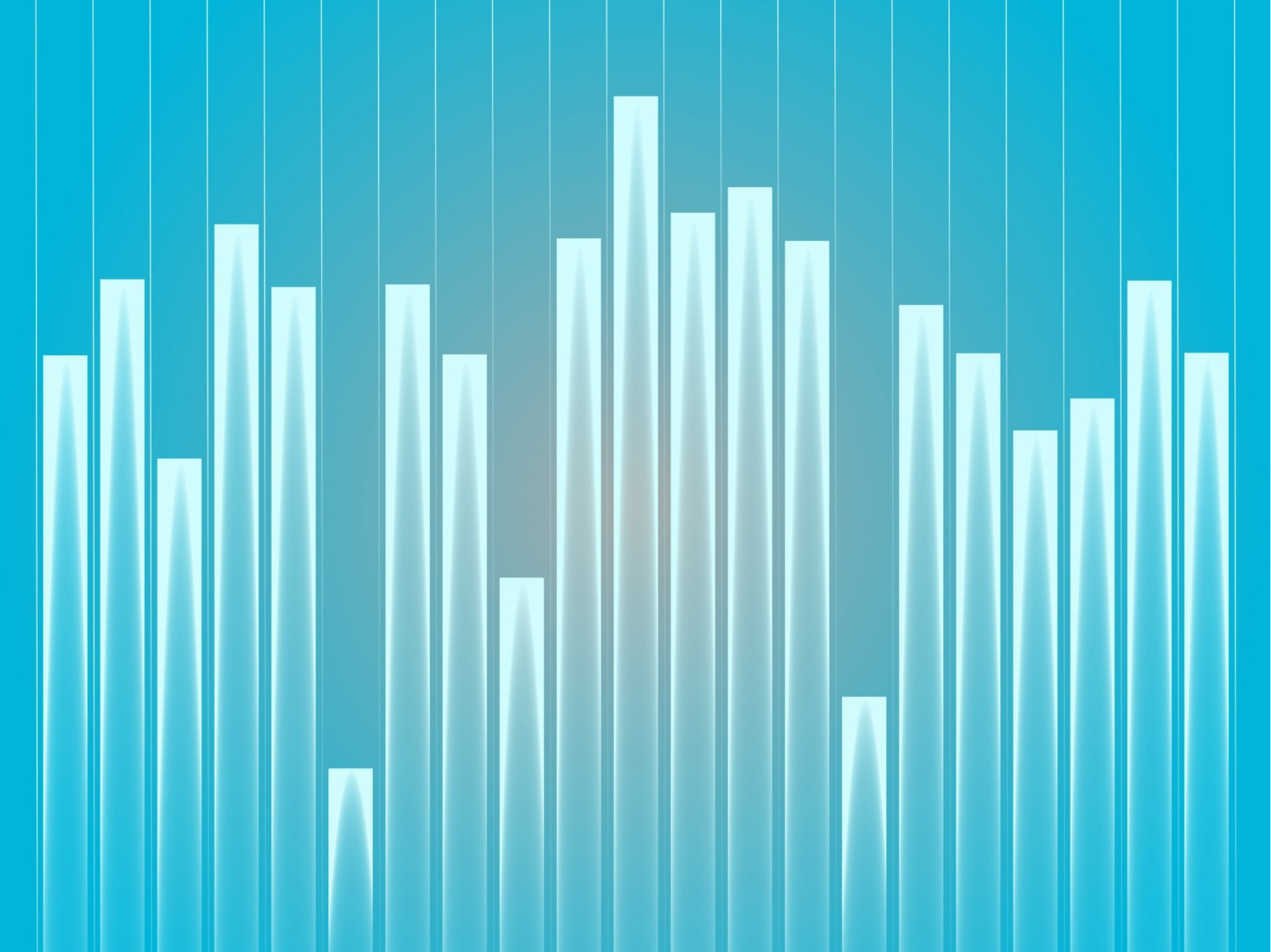
Referring_Page_URL..Please could you let me know when you have completed this removal or if you are unable to meet this request and the reasons for this.

Thank you in advance for your co-operation in this matter.

Kind Regards,
Webmaster
[www. \[redacted\] .au](http://www. [redacted] .au)

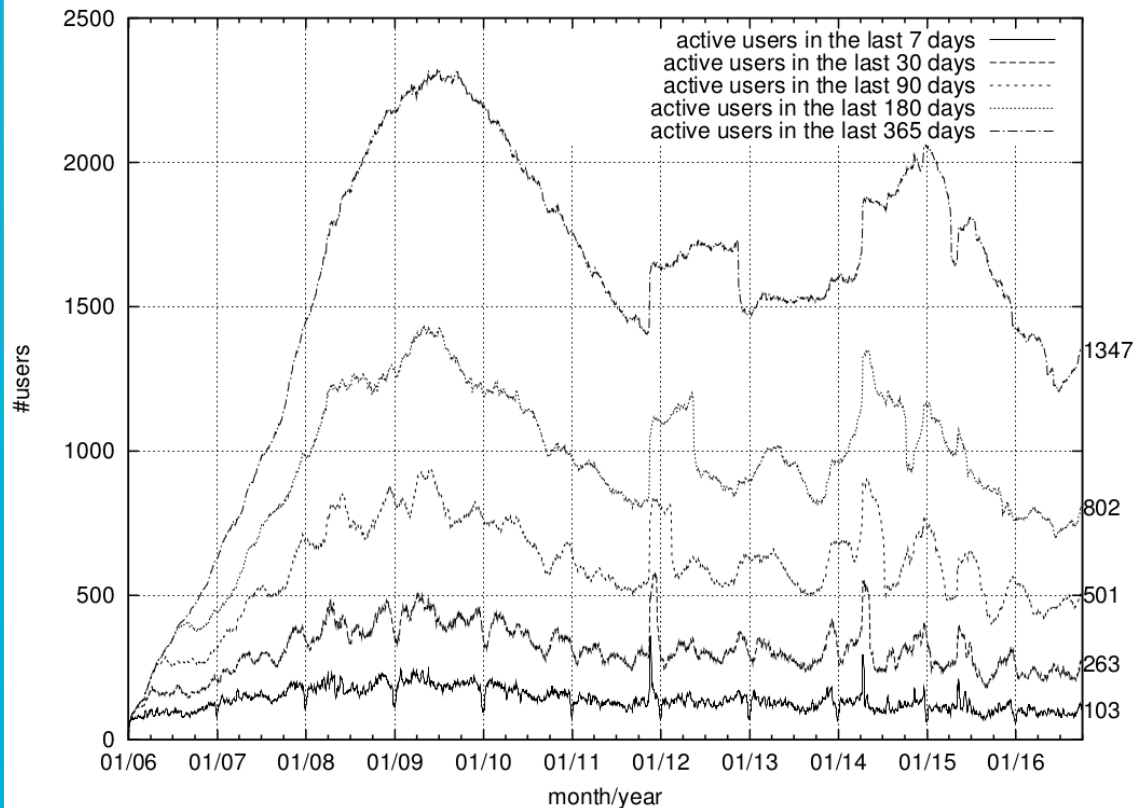


- be professional early on
- distribute responsibilities
- do not underestimate effort for user interaction
- be prepared for weird user behavior ;-)



	bookmarks	publications
resources:	625k	938k
posts:	727k	1045k
documents:	-	51k
users:	9k	10k

+ more than 5 million posts from DBLP & DNB
+ 2.5 million users that are flagged as spammers





- many more users would require more resources
- measures against spam crucial
- optimizing database schema + queries pays off



518811
ZEISS WETZLAR
GERMANY

A close-up, blue-tinted photograph of a microscope. The central focus is on the objective lens assembly, which has a silver-colored metal body. Engraved on the lens barrel are the numbers '518811', the brand name 'LEITZ WETZLAR', and 'GERMANY'. The background is blurred, showing other parts of the microscope and a slide on the stage.

Tagora

Semiotic Dynamics
in Online Social Communities

Info 2.0

Informational Self-Determination
in the Web 2.0



PUMA

Academic Publication Management

WISKIDZ

Changes in Academic Career Dynamics
in Germany

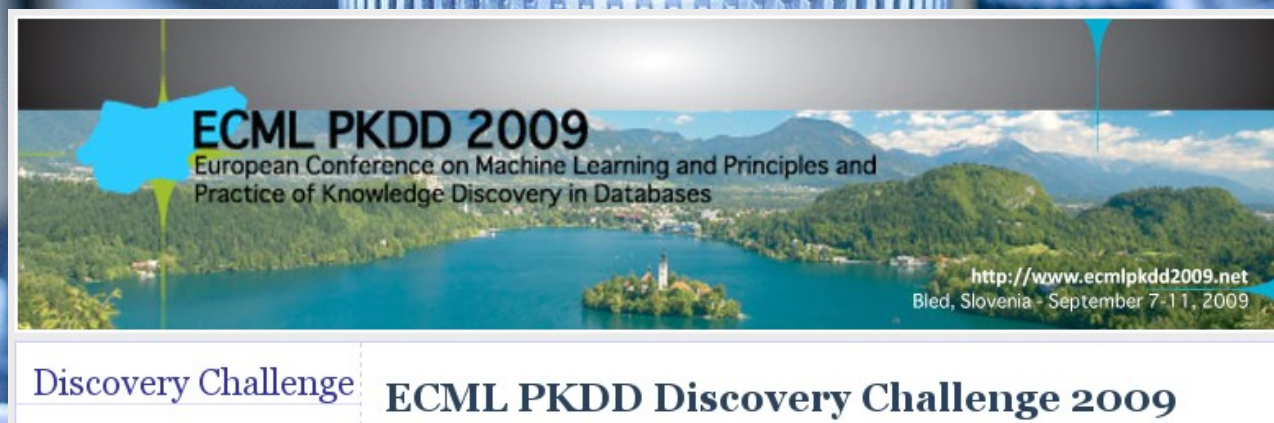


ECML 08 PKDD

RSDC'08 ECML PKDD Discovery Challenge 2008

15th. September 2008; Antwerp, Belgium

Discovery Challenge **ECML PKDD Discovery Challenge 2008**



ECML PKDD 2009
European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases

<http://www.ecmlpkdd2009.net>
Bled, Slovenia - September 7-11, 2009

Discovery Challenge **ECML PKDD Discovery Challenge 2009**

BibSonomy :: dumps for research purposes

The blue social bookmark and publication sharing system.

[About](#) [Datasets](#) [Description](#) [Log Data](#) [FAQ](#)

518811
ETZ WETZLAR
GERMANY



- unique opportunities if you run your own system
- can do A/B testing like the „big guys“
- research → production: a long and winding road
- need enthusiastic and persistent PhD students



PUMA

Akademisches Publikationsmanagement

PUMA ▾

Aktuelles

Lösungen ▾

Angebote

Szenarien ▾

Referenzen

Über uns ▾

Browser-Plugins

Firefox

Google Chrome

Discovery System

VuFind

eLearning

Moodle

OpenAccess

DSpace

PUMA-Publikationsmanagement

TYPO3

WordPress

Homepages und Institutswebseiten

Literatur-
verwaltung

Textver-
arbeitung

BibTeX, Endnote, Bibliographie-Stile

Converis

Forschungs-
informations-
system

<http://www.academic-puma.de/>



- 
- A close-up photograph of wooden Scrabble tiles on a wooden surface. The word "FUTURE" is spelled out in a row, with each letter tile showing its point value: F (4), U (1), T (1), U (1), R (1), and E (1). Several other letter tiles are scattered around, including S, Y, J, W, P, R, and E. A semi-transparent white box with a blue border is overlaid on the bottom left, containing a bulleted list.
- ORCID integration
 - web archiving
 - extended support for groups of users
 - federated search, fulltext document search
 - ...

A Cloud-Based Infrastructure for Community-Driven Data Mining on Research Outputs

jaeschke@L3S.de

