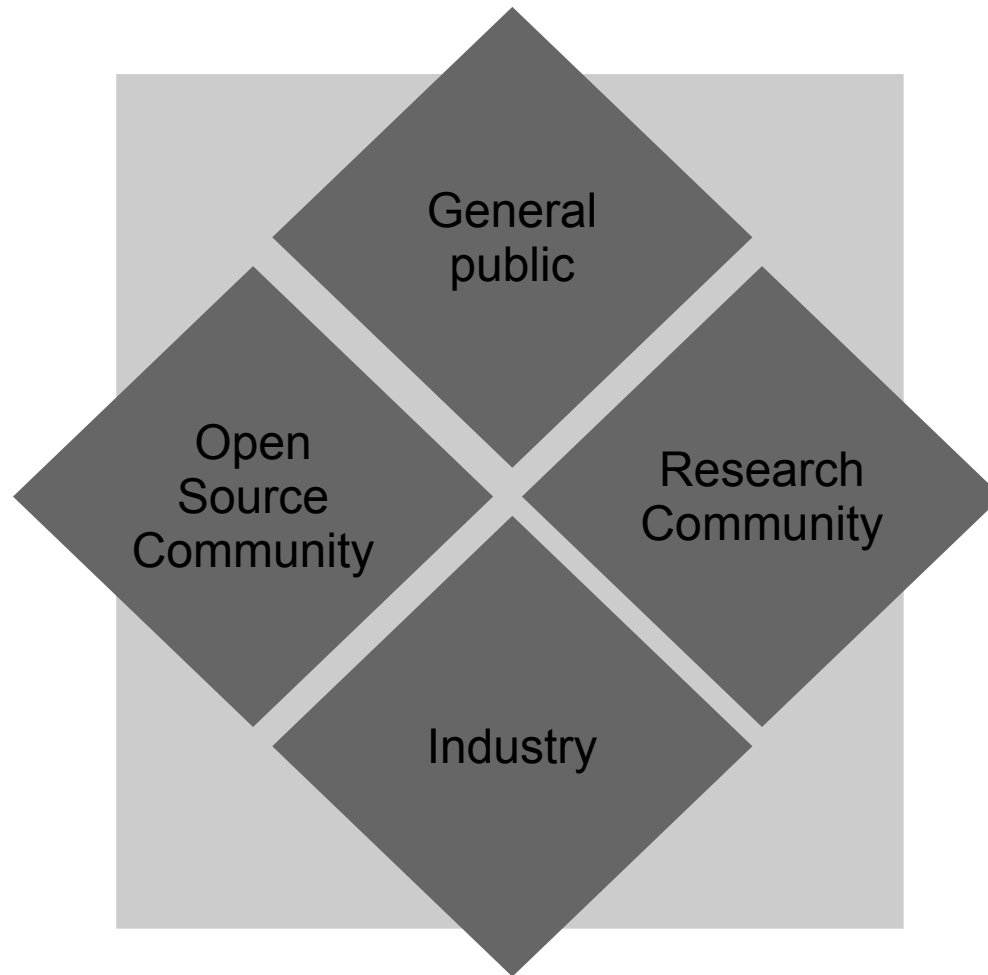




# Getting the word out

Foto thanks to <http://www.flickr.com/photos/photomic>

## Work package 7: Dissemination and Exploitation



Open Source Community

# Research Community



## Phase 1

Community bonding

## Phase 2

Get people to play

## Phase 3

Crowd sourcing

Goals of phase 1





Thanks to der\_rich for the fotos.









Actions in phase 1

✓ S	g Subject	From	Date
	Re: anybody want to set a record with Mahout?	Ted Dunning	10/02/25(Thu) 22:03
▽	Re: anybody want to set a record with Mahout?	Grant Ingersoll	10/02/25(Thu) 21:42
	Re: anybody want to set a record with Mahout?	Robin Anil	10/02/25(Thu) 21:44
▽	Re: anybody want to set a record with Mahout?	Jake Mannix	10/02/25(Thu) 22:38
	Re: anybody want to set a record with Mahout?	Ted Dunning	10/02/25(Thu) 22:46
	Re: anybody want to set a record with Mahout?	Jake Mannix	10/02/25(Thu) 21:39
▽	Proper way to dump kmeans clusters?	Drew Farris	10/02/26(Fri) 05:53
	Re: Proper way to dump kmeans clusters?	Robin Anil	10/02/26(Fri) 06:11
✖	▽ Re: Proper way to dump kmeans clusters?	cdwzwd	10/08/30(Mon) 11:52
✖	▽ Re: Proper way to dump kmeans clusters?	Jeff Eastman	10/08/30(Mon) 18:33
✖	Re: Proper way to dump kmeans clusters?	Joe Kumar	10/08/30(Mon) 20:36
▽	[OT] Interesting Stats	Robin Anil	10/02/26(Fri) 07:41
▽	Re: [OT] Interesting Stats	Eirik Bjørnsnes	10/02/26(Fri) 09:56
	Re: [OT] Interesting Stats	Robin Anil	10/02/26(Fri) 10:30
▽	[Jira] Created: (MAHOUT-311) Update assemblies to include c	Drew Farris (JIRA)	10/02/26(Fri) 14:40
	[Jira] Updated: (MAHOUT-311) Update assemblies to include	Drew Farris (JIRA)	10/02/26(Fri) 14:44
	[Jira] Updated: (MAHOUT-311) Update assemblies to include	Drew Farris (JIRA)	10/02/26(Fri) 14:46

**From:** Robin Anil <robin.anil@gmail.com>  
**To:** mahout-dev@lucene.apache.org  
**Subject:** Re: Proper way to dump kmeans clusters?  
**Date:** Fri, 26 Feb 2010 10:41:34 +0530  
**Reply-To:** mahout-dev@lucene.apache.org

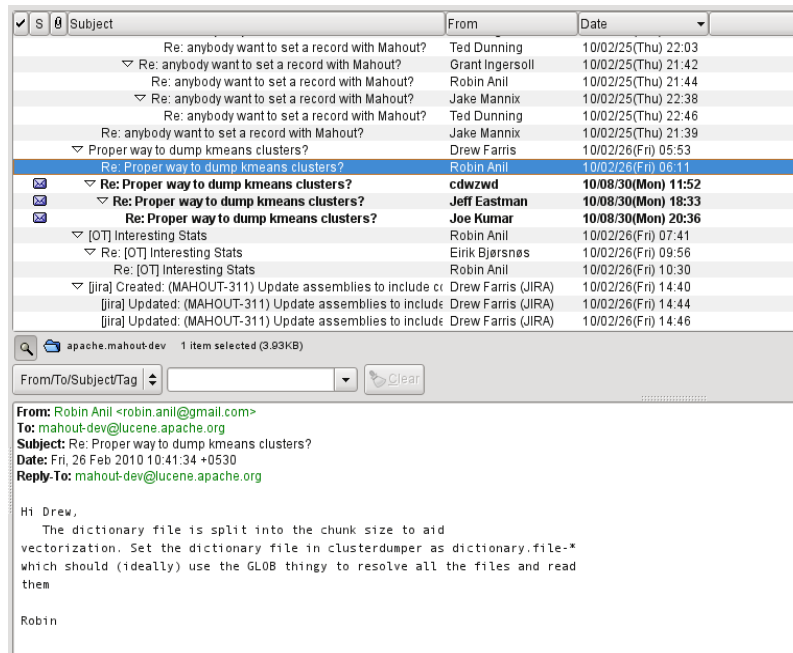
Hi Drew,

The dictionary file is split into the chunk size to aid vectorization. Set the dictionary file in clusterdumper as dictionary.file-\* which should (ideally) use the GLOB thingy to resolve all the files and read them

Robin

- Get in touch early – that is **now**.

# I want your slides!



- Get in touch early – that is **now**.



Image by: Patrick McEvoy



- Contribute back:
  - Patches.
  - Documentation.
  - Knowledge.
  - Data.
  - Patch reviews.
- = quite a bit of time.

Votes: 0

Watchers: [2](#)

#### Available Workflow Actions

☐ [Submit Patch](#)

☐ [Resolve Issue](#)

# Pre-condition for contributions

- Quality.
- Effectiveness:  
Use the source luke.
- Impact:  
Relevance to project's focus and user's needs.

citeulike  
sponsored by Springer



CiteULike MyCiteULike DICODE\_EU's CiteULike

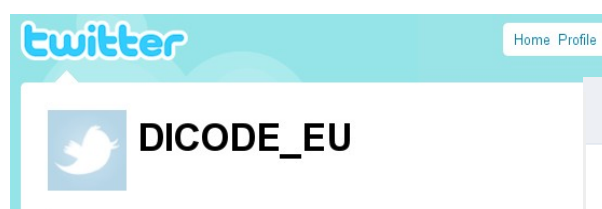
DICODE\_EU's library 13 articles

Search Watch Connect Export Sort Hide Details

**Detecting the origin of text segments efficiently**  
In WWW '09: Proceedings of the 18th international conference on World wide web (2009), pp. 61-72  
by Ossama A. Hamid, Behshad Behzadi, Stefan Christoph, Monika Henzinger  
posted to [google crawler](#) by [DICODE\\_EU](#) on 2010-08-05 12:55:11 \*\* [along with 3 people](#)  
■ Abstract ■ My Copy

**Dremel: Interactive Analysis of Web-Scale Datasets**  
In The 36th International Conference on Very Large Data Bases, Vol. 3 (September 2010)  
by Sergey Melnik, Andrey Gubarev, Jing J. Long, et al.  
posted to [mapreduce](#) [google dremel](#) by [DICODE\\_EU](#) on 2010-08-05 12:52:59 \*\* [along with 1 person](#)  
■ Copy

**IRLbot: scaling to 6 billion pages and beyond**  
In WWW '08: Proceedings of the 17th international conference on World Wide Web (2008), pp. 497-506



DICODE

Wall Info Discussions Photos Video Events

Write something...

Attach: [image icon] [document icon] [video icon] [audio icon]

**Dicode Eu** DICODE\_EU: Now on to putting some content into our project web page.

Read: Full story  
Source: Twitter URL  
Published: 2010-08-24 15:59:07 GMT  
August 24 at 6:06pm via RSS Graffiti · Comment · Like · Share

**DICODE**

GOALS OPEN SOURCE STRATEGY PROJECTS

**Example Use Cases**

Tue, 08/31/2010 - 12:30 — Isabel

The achievement of the project's objectives will be validated through carefully selected to:

- Address clearly established problems, widely recognised in the community
- Cover the full range of features and functionalities of the project, including collaboration and decision making paradigms (that may be domain-specific), and dealing with various types of large scale and heterogeneous sources.

A short description of each use case is given below.

# • Be available and visible:

- Online and
- Offline.



**DEVOXX**  
the pure community conference

November 15th till 19th 2010 - Metropolis Antwerp, Belgium

Home Javaholics Schedule Partners News Practical

University Tools in Action Conference BOFs Quickies

**ApacheCon**  
NORTH AMERICA 2010

1-5 November 2010 | [Westin Peachtree, Atlanta](#)



Upcoming

28th International Conference on Machine Learning (ICML 2011)

will be in Seattle, Washington, USA, date to be determined



- NIPS Home
- NIPS 2010
- Conference Info
- My Account
- Previous Meetings
- Memorial Fund
- Publications
- Sponsors / Exhibitors
- Directory

NIPS 2011

Tutorials – December 12, 2011  
Granada Congress and Exhibition Centre, Granada, Spain

Conference Sessions – December 13-15, 2011  
Granada Congress and Exhibition Centre, Grenada, Spain

Workshops – December 16-17, 2011  
Melia Sierra Nevada & Melia Sol y Nieve, Sierra Nevada, Spain



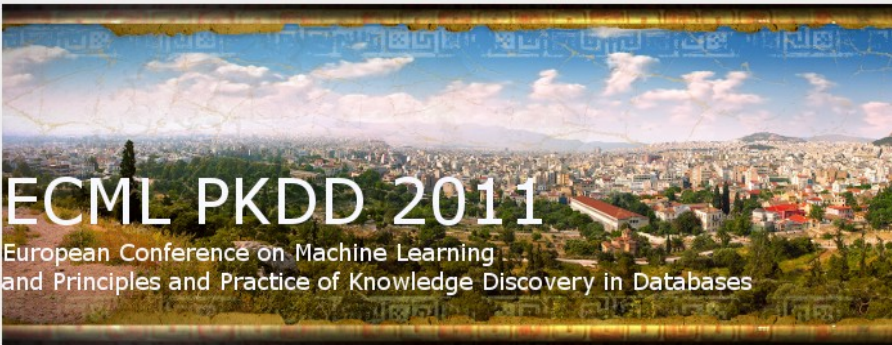
ACM SIGIR 2011

July 24-28, 2011  
The 34<sup>th</sup> Annual  
ACM SIGIR Conference



KDD Conferences

Next Year's Co



- WELCOME
- INFORMATION
- Key Dates

WELCOME

The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD) will take place in Athens, Greece from September 12th to 16th, 2011.

Goals of phase 2











Actions in phase 2







ACM Special Interest Group on Knowledge Discovery and Data Mining

KDD Cup

[Home Page](#)  
[KDD Cup 2008](#)  
[KDD Cup 2007](#)


### KDD Cup Center

KDD Cup is the annual Data Mining and Knowledge Discovery competition organized by [ACM Special Interest Group on Knowledge Discovery and Data Mining](#), the leading professional organization of data miners.

Below are links to the descriptions of all past tasks.

**LEARNING TO RANK CHALLENGE**
from YAHOO! LABS

[Home](#)
[Datasets](#)
[Instructions](#)
[Registration](#)
[Submission](#)
[Leaderboard](#)
[FAQs](#)
[Workshop](#)



# CONGRATULATIONS!

## LEARNING TO RANK CHALLENGE



## Pascal Large Scale Learning Challenge

[About](#)
[Instructions](#)
[Registration](#)
[Su](#)

ration



Goals of phase 3





<http://www.flickr.com/photos/riacale/4126971310/>



Actions for phase 3









## Machine Learning Open Source Software

[Home Page](#)

[Papers](#)

[Submissions](#)

[News](#)

[Scope](#)

[Editorial Board](#)

[Announcements](#)

[Proceedings](#)

To support the open source software movement, JMLR MLOSS publishes contributions related to toolboxes or even languages for scientific computing. Submission instructions are available [here](#).

A Library for Locally Weighted Projection Regression

*Stefan Klanke, Sethu Vijayakumar, Stefan Schaal*, 9(Apr):623--626, 2008.

[\[abs\]](#)[\[pdf\]](#)   [\[code\]](#)[\[mloss.org\]](#)

Shark

*Christian Igel, Verena Heidrich-Meisner, Tobias Glasmachers*, 9(Jun):993--996, 2

[\[abs\]](#)[\[pdf\]](#)   [\[code\]](#)[\[mloss.org\]](#)

LIBLINEAR: A Library for Large Linear Classification

*Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, Chih-Jen Lin*,

[\[abs\]](#)[\[pdf\]](#)   [\[code\]](#)[\[mloss.org\]](#)



Prerequisites



Release early and often.

Events	
•	Wed 9/15/2010: Mahout@JUGBB
•	Mon 11/1/2010: Apache Con NA
•	Mon 11/15/2010: Devvxx

Release early and often.

Release as in online and **publicly available**.

Early as in until **end of October**.

Often as in **every month** a new version.











