

# Open source is just about the source, isn't it?

“The Apache Way – 19 years of OSS experience in 40 minutes”



Based on a talk I first gave at FrOSCon 2016, Apache Way talks by Lars Eilebrecht, Justin Ehrenkrantz, Brett Porter, Nick Burch, Rich Bowen, feedback from Shane Curcuru, input from Bertrand Delacretaz, Pieter Hintjens, Zaheda Bhorat, corrections from Sally Khudairi, Mark Thomas, Phil Steitz, Daniel Ruggeri, stuff I read in “Producing Open Source Software” by Karl Fogel, stuff I read in “Building successful online communities” by Kraut/ Resnick, stuff I read in “Social Architecture” by Pieter Hintjens, and numerous conversations on Apache mailing lists.

And most likely many I forgot to mention above.



# **Isabel Drost-Fromm**



**Open Source Strategist Europace AG**  
<http://www.europace.de>

**(Board) Member Apache Software Foundation**  
<http://www.apache.org>

**Co-founder Apache Mahout**  
<http://mahout.apache.org>

**Co-founder Berlin Buzzwords**  
<http://www.berlinbuzzwords.de>

**Co-founder FOSS Backstage**  
<http://www.foss-backstage.de>



why  
you even care

Image taken shortly before FOSS Backstage Microsummit Berlin @ Europace AG  
Batches were printed based on the names ppl used for registration.

Mission: Provide **software for the public good** by providing services and **support** for many **like-minded software project communities of individuals** who choose to join the ASF.

Funding: Individual donations + corporate sponsorships.

US 501(c)(3) nonprofit charitable organization

Established in 1999.

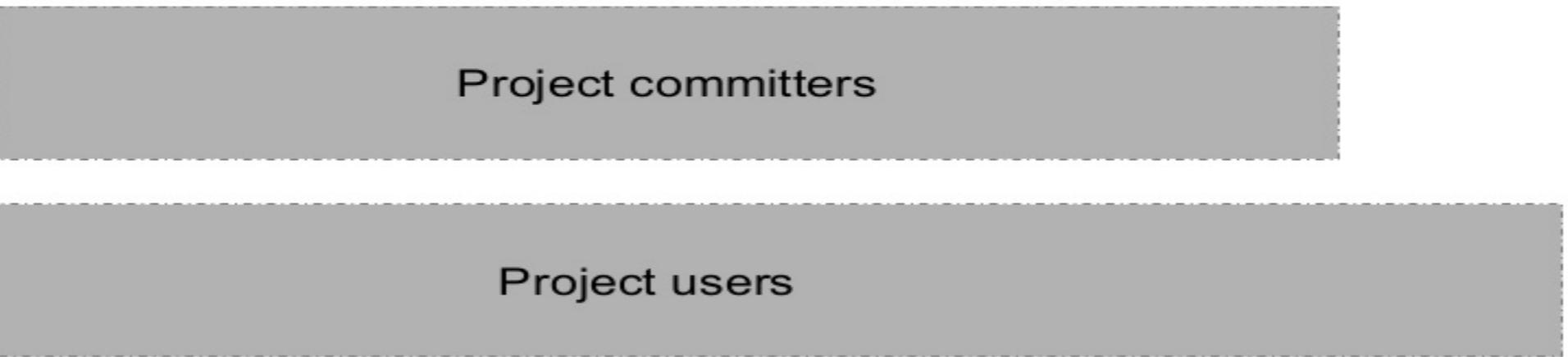


Project users

*Yonik's law of Patches: A half-baked patch in Jira, with no documentation, no tests and no backwards compatibility is better than no patch at all.*

<https://wiki.apache.org/solr/HowToContribute>

Project users



Project committers

Project users



Image based on graphic by Brett Porter.

Apache Software Foundation  
Members

Project Management Committees

Project committers

Project users

```
graph TD; A[Apache Software Foundation Members] --- B[Board of Directors]; B --- C[Project Management Committees]; C --- D[Project committers]; D --- E[Project users]
```

Apache Software Foundation  
Members

Board of Directors

Project Management Committees

Project committers

Project users

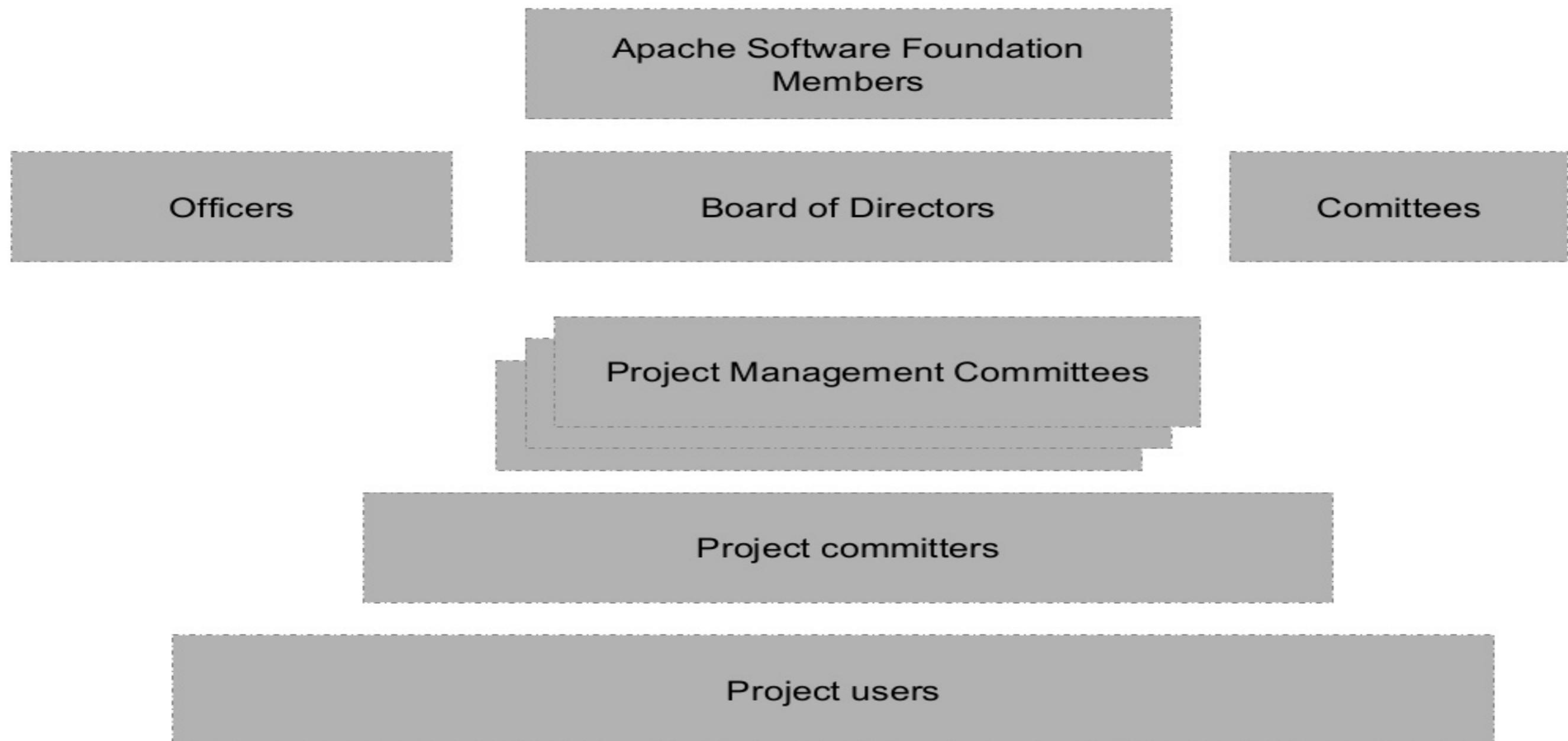


Image based on graphic by Brett Porter.

# Founding fathers



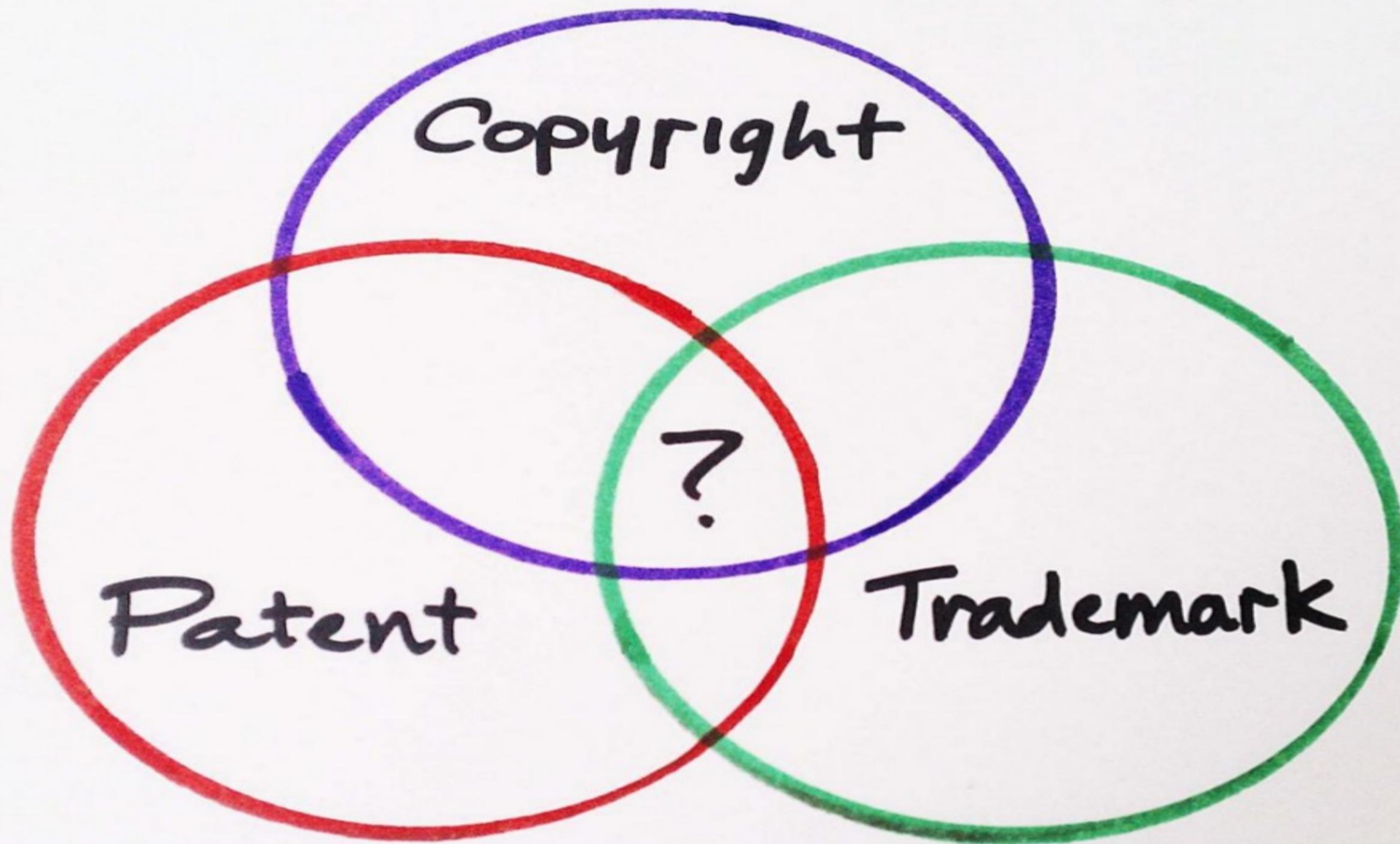
Image taken at Open Source Summit 2017, Prague; talk by Zaheda Bhorat,  
<https://osseu17.sched.com/event/ByN2/love-what-you-do-everyday-zaheda-bhorat-aws>.

[https://blogs.apache.org/foundation/entry/success\\_at\\_apache\\_project\\_independence](https://blogs.apache.org/foundation/entry/success_at_apache_project_independence)



Image "Swiss Flag" by crackers93

<https://www.flickr.com/photos/crackers93/2832784903> (CC-By-2.0)





Copyright

I care about any and all of my downstream users to have all of “use”, “study”, “share”, “improve”:

All I want to ensure is that my very own project gives the “use”, “study”, “share”, “improve” freedoms:

## Copyleft OSS

LGPL for libraries, especially if there are other similar libraries

AGPL for server software

GPL for everything else

## Non-Copyleft OSS

Small enough so you don't care

Libraries to push standards forward

Projects to change established economics.

All I want to ensure is that my very own project gives  
the “use”, “study”, “share”, “improve” freedoms:

### Non-Copyleft OSS

Apache License 2.0

<https://www.apache.org/licenses/LICENSE-2.0>



Image "Copyright, Patent, or Trademark" by BusinessSarah  
<https://www.flickr.com/photos/businesssarah/5977958263> (CC-BY-2.0)

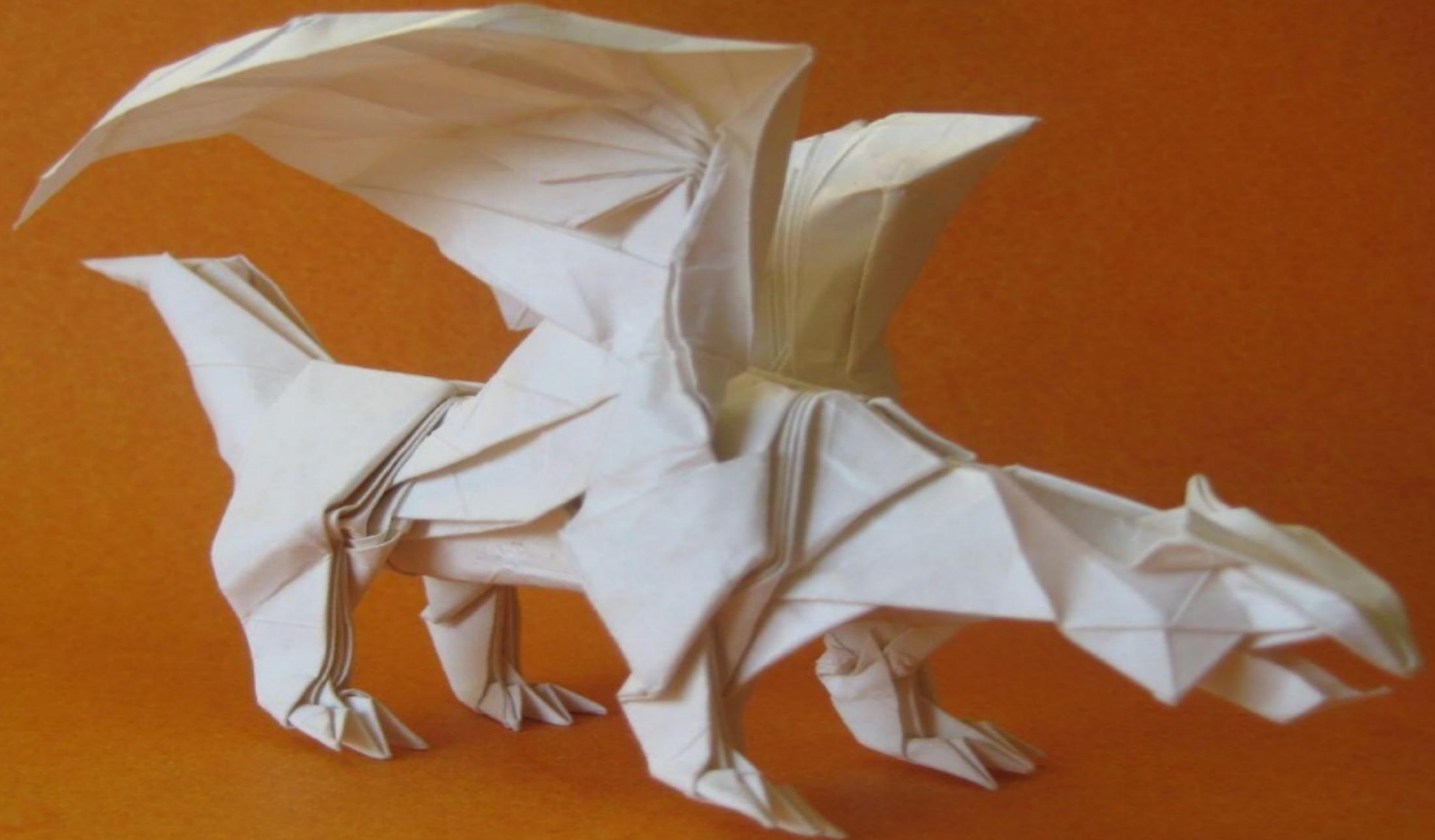


Image "Dragon" by Joseph Wu

<https://www.flickr.com/photos/josephwuorigami/1367278646> (CC BY-NC-ND 2.0)

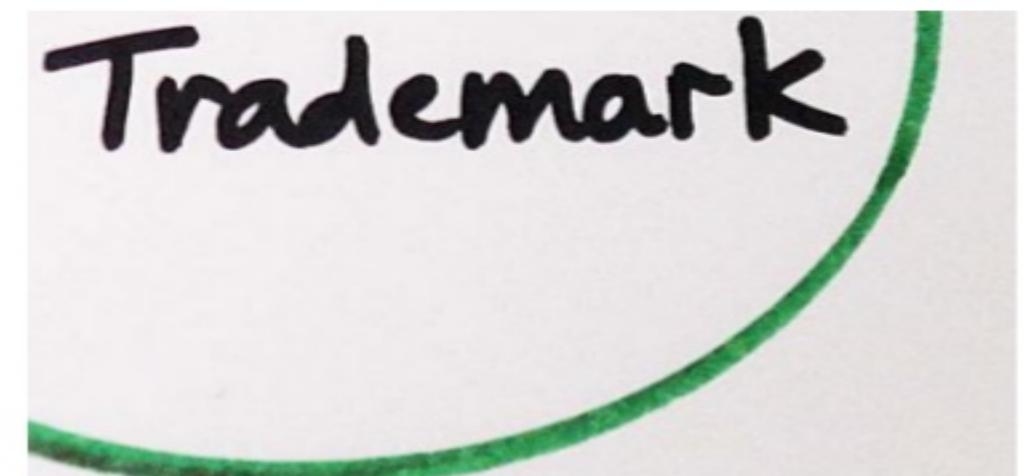




Photo by Daniele Levis Pelusi on Unsplash



Image taken at Berlin Buzzwords.

# Apache: Community over code.

“A project without people is a dead project (or at least deep asleep).”

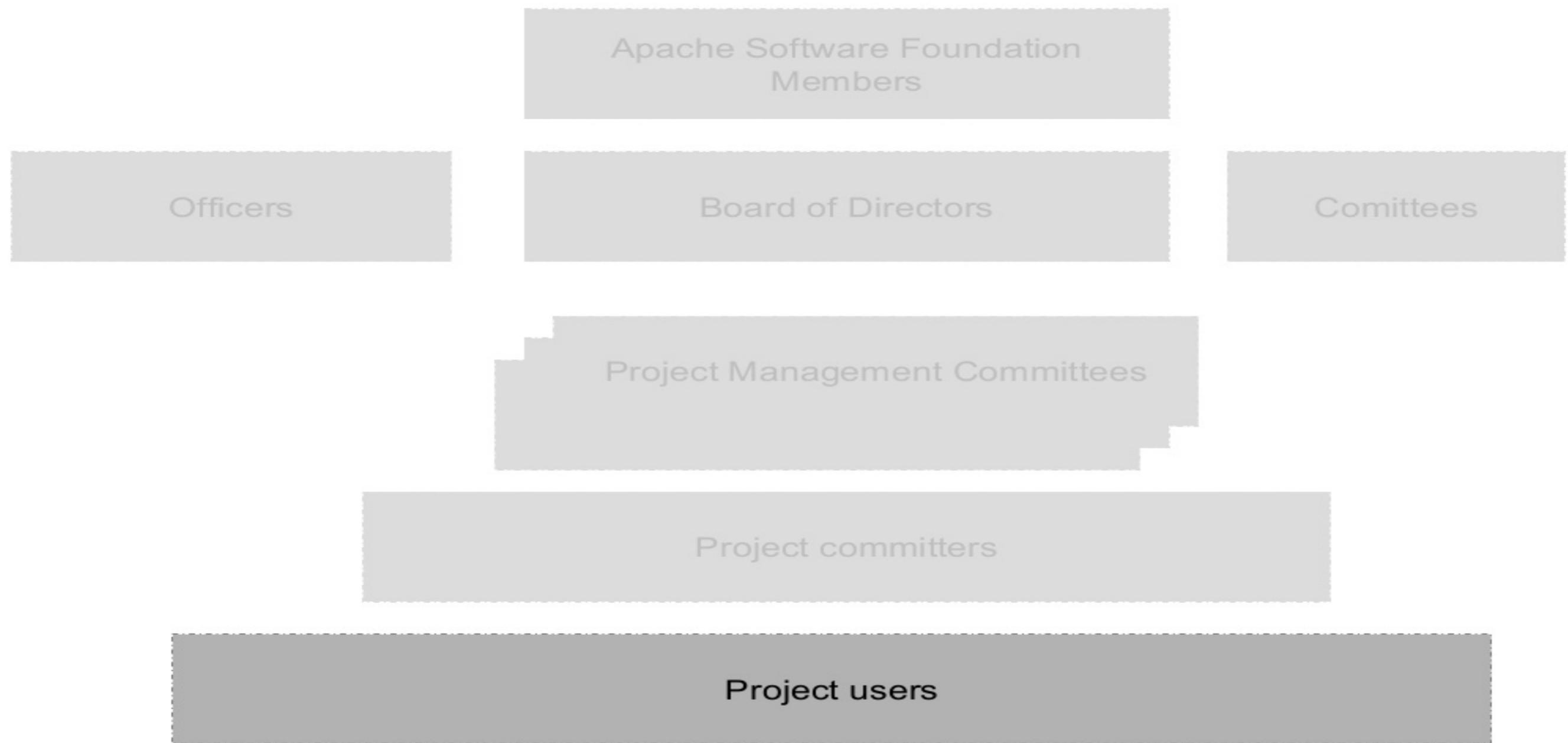


Image based on graphic by Brett Porter.

<https://twitter.com/ApacheMahout>

Startseite Mitteilungen Nachrichten

Tweets 1.247 Folge 36

**Apache Mahout**  
@ApacheMahout Folgt dir

Apache Mahout is an open source Apache Foundation project for scalable machine learning.

[mahout.apache.org](http://mahout.apache.org)

Beigetreten Januar 2013

[Tweet an](#) [Nachricht](#)

43 Follower, die du kennst

Fotos und Videos

**Tweets** Tw

 Apache Mahout Chapter is published  
Original



 Apache Mahout Trevor Gramineus according aspiring #BigData

Secure | <https://stackoverflow.com/questions/tagged/mahout>

stack overflow Questions Developer Jobs Tags Use

The results are in! See the 2018 Developer Survey results. »

## Tagged Questions

info

newest

Apache Mahout open source scalable machine learning project

[learn more...](#) [top users](#) [synonyms](#)

-2

votes

### Content based recommender system using mahout

I know that mahout implement collaborative filtering approach. I need to calculate similarity between 2 products

Secure | <https://www.google.com/alerts#2:0>

# Alerts

Monitor the web for interesting news

← Apache mahout

HOME

**jaxenter**  
Portal für Java™, Enterprise Architekturen, SOA

Java EE

republik Java Magazin Eclipse Magazin W-JAX 2012 Autor werden!

Oktöber 2012 | Artikel

## Apache Lucene 4.0

Neue Features im Überblick

Text: Uwe Schindler



Apache Lucene geht in die nächste Runde: Lange Datentypen und Strukturen zur Ablage des Index, die Version vor 13 Jahren vorstellt, wurde dieses Form langjähriger Entwicklungszeit erscheint nun parallel 4.0. Neben dem Update des Indexformats auf ein gibt es aber auch noch viele andere Neuerungen: Indexieren mit abgelegt werden können, ist es auch von Lucene durch moderne Varianten wie BM25 zu von neuen Codecs und Ranking-Funktionen ausge-



TM



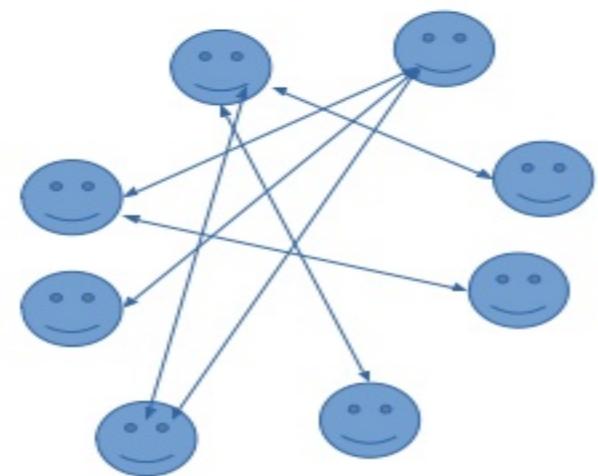
Image by NS Newsflash

<http://www.flickr.com/photos/62693815@N03/6277336776/>



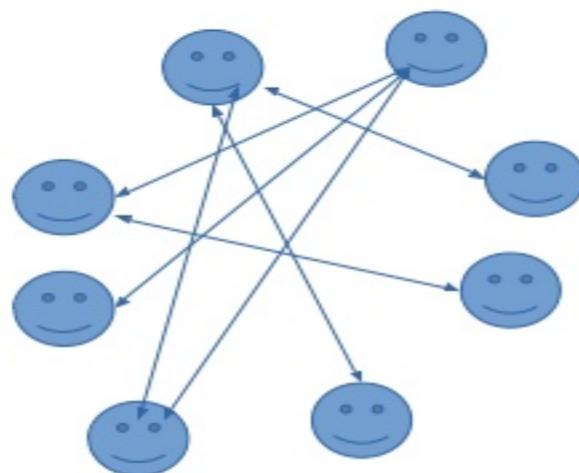
Image taken earlier this year at FrOSCon.  
Shirts available at FSFE booth e.g. at FOSDEM or Chemnitzer Linuxtag.

## Mess Media



# Central Hubs

## Mess Media



The top screenshot shows a GitHub repository page for 'apache / mahout'. It displays a list of pull requests, with one prominent entry: 'MAHOUT-2019 Sparse speedup' by 'Hoai Nguyen' opened 8 days ago. The middle screenshot is a screenshot of a Pony Mail inbox for 'user@mahout.apache.org'. It shows two recent messages: one from 'Hoa Nguyen' about running Mahout on a Spark cluster, and another from 'n...@apache.org' about scanning data. The bottom screenshot is a JIRA issue tracking interface for 'Mahout / MAHOUT-2000'. It shows a list of open issues and a detailed view of 'MAHOUT-2000: Automatic Probing for GPU/OMP modules at the creation of a MahoutDistributedContext'. The issue is categorized as a bug, has a priority of critical, and is currently unresolved.

Thanks to Bertrand Delacretaz for inspiration.

If it didn't happen on the mailing list,  
it didn't happen.



If it didn't happen on the mailing list,  
it didn't happen.

**user@**

**dev@**

commits@

issues@

private@



# http://lists.apache.org

Secure | https://lists.apache.org/list.html?dev@spark.apache.org

Apache Mail Archives Search this list Less than 1 month ago Type search terms... Advanced reviews@ issues@ commits@ user@ dev@ Available projects: Show next

Archive: Stats for this blob of emails:  
211 emails sent by 90 people, divided into 53 topics. [details]

View mode: threaded

### dev@spark.apache.org, last month

**José Raúl Pérez Ro... cache OS memory and spark usage of it**  
Hi, When I issue a "free -m" command in a host, I see a lot of memory used for cache in OS, however Spark Streaming is not able to request that memory for its ...  
2 people 1 reply 2018-04-11 06:57

**Louis Hust How to use disk instead of just InMemoryRelation when use JDBC datasource in SPARKSQL?**  
We want to extract data from mysql, and calculate in sparksql. The sql explain like below. == Parsed Logical Plan == > 'Sort [revenue DESC NULLS LAST], true ...  
0 replies 2018-04-11 03:43

**Tushar Singhal Accessing Hive Tables in Spark**  
Hi Everyone, I was accessing Hive Tables in Spark SQL using Scala submitted by spark-submit command. When I ran in cluster mode then got error like : Table n...  
3 people 2 replies 2018-04-10 06:58

**Niu Zhaojie pyspark.daemon exhaust a lot of memory**  
Hi All, We are running spark 2.1.1 on Hadoop YARN 2.6.5. We found the pyspark.daemon process consume more than 300GB memory. However, according to ht...  
0 replies 2018-04-10 04:35

**Nicholas Chammas Correlated subqueries in the DataFrame API**  
I just submitted SPARK-23945 <<https://issues.apache.org/jira/browse/SPARK-23945>> but wanted to double check here to make sure I didn't miss something fund...  
0 replies 2018-04-09 22:56

**Matthias Boehm Re: Fair scheduler pool leak**  
well, the point was "in a programmatic way without the need for additional configuration files which is a hassle for a library" - anyway, I appreciate your comments....  
2 people 1 reply 2018-04-09 16:28

**Li Jin Clarify window behavior in Spark SQL**  
Hi Devs, I am seeing some behavior with window functions that is a bit unintuitive and would like to get some clarification. When using aggregation function with ...  
4 people 8 replies 2018-04-09 12:06



Photo by Ricardo Gomez Angel on Unsplash



Image "Help Me. VTda.info" by vtdainfo  
<https://flic.kr/p/7rrxxh> (CC-BY 2.0)



Photo by Thiago Victal on Unsplash

+1



-1



**Those who do the work take the decisions.**

<https://www.apache.org/foundation/how-it-works.html#decision-making>™



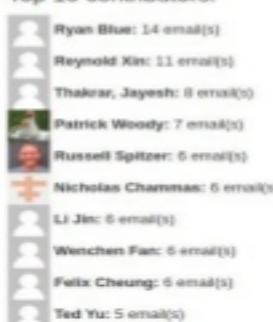
# Apache Ponymail

<http://ponymail.apache.org>

← → ⌂ | Secure | <https://lists.apache.org/list.html?dev@spark.apache.org>

Apache Mail Archives  Type search terms...  reviews@ issues@ commits@ user@ dev@ Available projects:

Archive:       Stats for this blob of emails: 211 emails sent by 90 people, divided into 53 topics [details] 

Top 10 contributors:   
Ryan Blue: 14 email(s)  
Reynold Xin: 11 email(s)  
Thakrar, Jayesh: 11 email(s)  
Patrick Woody: 7 email(s)  
Russell Spitzer: 6 email(s)  
Nicholas Chammas: 6 email(s)  
Li Jin: 6 email(s)  
Wenchen Fan: 6 email(s)  
Felix Cheung: 6 email(s)  
Ted Yu: 5 email(s)

View mode:

dev@spark.apache.org, last month

**José Raúl Pérez Ro... cache OS memory and spark usage of it**   2018-04-11 06:57  
Hi, When I issue a "free -m" command in a host, I see a lot of memory used for cache in OS, however Spark Streaming is not able to request that memory for its ...

**Louis Hust How to use disk instead of just InMemoryRelation when use JDBC datasource in SPARKSQL?**  2018-04-11 07:43  
We want to extract data from mysql, and calculate in sparksql. The sql explain like below. == Parsed Logical Plan == > 'Sort [revenue DESC NULLS LAST], true ...

**Tushar Singhal Accessing Hive Tables in Spark**   2018-04-10 04:58  
Hi Everyone, I was accessing Hive Tables in Spark SQL using Scala submitted by spark-submit command. When I ran in cluster mode then got error like - Table n...

**Niu Zhaojie pyspark.daemon exhaust a lot of memory**  2018-04-10 04:25  
Hi All, We are running spark 2.1.1 on Hadoop YARN 2.6.5. We found the pyspark.daemon process consume more than 300GB memory. However, according to ht...

**Nicholas Chammas Correlated subqueries in the DataFrame API**  2018-04-09 22:56  
I just submitted SPARK-23945 <<https://issues.apache.org/jira/browse/SPARK-23945>> but wanted to double check here to make sure I didn't miss something fund...

**Matthias Boehm Re: Fair scheduler pool leak**   2018-04-09 14:28  
well, the point was "in a programmatic way without the need for additional configuration files which is a hassle for a library" - anyway, I appreciate your comments....

**Li Jin Clarify window behavior in Spark SQL**   2018-04-09 12:06  
Hi Devs, I am seeing some behavior with window functions that is a bit unintuitive and would like to get some clarification. When using aggregation function with ...

**Open**

bwebster wants to merge 1 commit into apache:master from bwebster:master

Conversation 0

Commits 1

Files changed 2

+6 -3

Showing 2 changed files with 6 additions and 3 deletions.

Unified Split

mrlegacy/src/main/java/org/apache/mahout/vectorizer/HighDFWordsPruner.java

View

		@@ -79,8 +79,7 @@ private static void pruneVectorsPartial(Path input, Path output, Path dictionary
79	79	+ "org.apache.hadoop.io.serializer.WritableSerialization");
80	80	conf.setLong(MAX_DF, maxDF);
81	81	conf.setLong(MIN_DF, minDF);
82		- DistributedCache.setCacheFiles(
83		- new URI[]{dictionaryFilePath.toUri()}, conf);
84	82	+ DistributedCache.addCacheFile(dictionaryFilePath.toUri(), conf);
85	83	Job job = HadoopUtil.prepareJob(input, output, SequenceFileInputFormat.class,
86	84	Mapper.class, null, null, WordsPrunerReducer.class,

mrlegacy/src/main/java/org/apache/mahout/vectorizer/pruner/WordsPrunerReducer.java

View

		@@ -16,7 +16,9 @@
16	16	* limitations under the License.
17	17	*/
18	18	
19	19	+import java.net.URI;
20	20	import org.apache.hadoop.conf.Configuration;
21	21	+import org.apache.hadoop.filecache.DistributedCache;
22	22	import org.apache.hadoop.fs.Path;
23	23	import org.apache.hadoop.io.IntWritable;
24	24	import org.apache.hadoop.io.LongWritable;



Image "Stopwatch" by Search Engine People Blog  
<https://www.flickr.com/photos/sepblog/4072462666> (CC-By 2.0)

Finalizar su compra es muy sencillo. Su pedido está pendiente de pago. Ahora puede comprobar que los datos de entrega son correctos y elegir el método de pago. Si desea comprobar el contenido del pedido, puede hacerlo en cualquier momento en la opción **sus pedidos** del menú superior.

1. Compruebe sus datos de entrega Si la dirección de entrega no es correcta [pulse aquí para editarla](#)

- 1:     Nombre &nbsp;<%= Apellidos %></strong>
- 2:     Apellidos &nbsp;<%= url\_main\_secure %>icons/user.gif);"><strong>&nbsp;<%= url\_main\_secure %>icons/id.gif);"><strong>&nbsp;<%= url\_main\_secure %>icons/book.gif);">
- 3:     <% if direccion2<>"" and cp2<>"" then %>
- 4:         <%= Direccion2 %><br>
- 5:         <%= CP2 %>&nbsp;<%= Localidad2 %><br>
- 6:         <%= Provincia2 %><br>
- 7:         <%= Pais %><br>
- 8:         <% else %>
- 9:             <%= Direccion %><br>
- 10:          <%= CP %>&nbsp;<%= Localidad %><br>
- 11:          <%= Provincia %><br>
- 12:          <%= Pais %><br>
- 13:          <% end if %>
- 14:     </li>
- 15:     <li style="list-style-type: none;">
- 16:         if movil<>"" then response.write("&nbsp;/&nbsp; " & Movil) %></li>
- 17:         <li style="list-style-type: none;">
- 18:             response.write("&nbsp;/&nbsp; " & Email) %></li>
- 19:     </ul>
- 20:     <div class="spacer"></div>
- 21:     <h2>2. Elija la forma de pago</h2>

Image by hercios

<http://www.flickr.com/photos/hercios/3022551609>

"Reviewing code" (CC-By-NC-SA)





Image by \_Faraz

<http://www.flickr.com/photos/faraz/332505613/>



Image by ryonedwards

<http://www.flickr.com/photos/ryonedwards/4128380265>



Image by rogercarr

<http://www.flickr.com/photos/rogercarr/3667924285/>



Image by viZZual.com

<http://www.flickr.com/photos/vizzual-dot-com/2888418520>

"Thank you card" (CC-BY-NC-SA 2.0)

"ThankYou\_rush's" (CC-BY-SA 2.0)

"merci" (CC-BY 2.0)

"Thank You Note" (CC-BY-NC-SA 2.0)

Contributions include **not only source code**, but also documentation, constructive bug reports, constructive discussions, marketing and generally anything that adds value to the project.



**Merit does not go away.**

<https://www.apache.org/dev/committers.html#committer-set-term>





Image "Money" (CC-BY 2.0) by Tax Credits  
<https://www.flickr.com/photos/76657755@N04/7027604401>



Image "Hats" (CC-By-N  
by arbyreed <https://flic.>

**At Apache you participate as an individual.**

<https://community.apache.org/projectIndependence.html>



**At Apache you pay with time and love.**



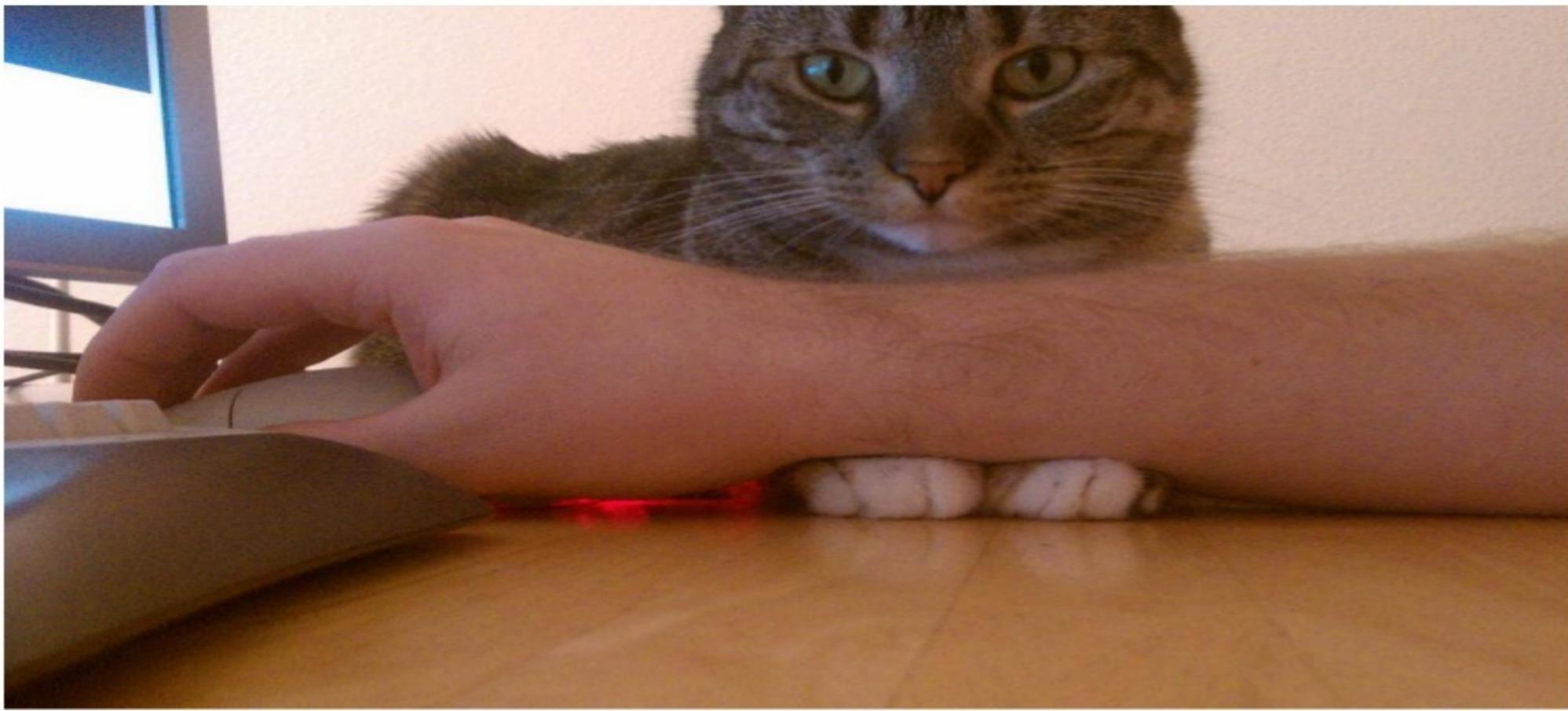


Image "Doxy cares about my ergonomics" (CC-By-SA 2.0) by Abraham Williams  
<https://www.flickr.com/photos/4braham/5153162831>

<https://s.apache.org/volunteeritis>

<https://s.apache.org/youwillburnout>



Image "Flying Pigeon bakfiets loaded up for the bike move" by Umberto Brayj  
<https://www.flickr.com/photos/ubrayj02/6917443031> (CC-BY 2.0)

**At Apache you pay with time and love ...**

**mostly.**





Machines to host infrastructure.

People to take care of these machines.

Press, Legal, Admin, Travel support, Trademarks.



Google Custom



The Apache Way

Contribute

ASF Sponsors

## SUPPORT APACHE

### MAKE A DONATION TODAY:

<a href="#">&gt; with a debit/credit card</a>	<a href="#">&gt; with PayPal</a>	<a href="#">&gt; mail in a check</a>	<a href="#">&gt; using ACH bank payment</a>
<a href="#">&gt; become an ASF Sponsor</a>	<a href="#">&gt; with Amazon</a>	<a href="#">&gt; using Bitcoin</a>	<a href="#">&gt; shop smile.amazon.com</a>
<a href="#">&gt; donate your car</a>	<a href="#">&gt; using Corporate Gifting</a>		

You're not alone.

[dev@community.apache.org](mailto:dev@community.apache.org)

[members@apache.org](mailto:members@apache.org)

Use ApacheCon, come to Apache  
Roadshow, join FOSS Backstage.





Image "Fire Hose by Matthew Heun" (CC-By-NC-SA 2.0) by Derrick Story  
<https://www.flickr.com/photos/thedigitalstory/5106832748/>



Image "Poison" (CC-By-NC-SA 2.0) by Broken Clouds  
<https://www.flickr.com/photos/34814824@N00/2805079971>



Image "frocon\_10\_IMG\_9710" (CC-By-NC 2.0) by saschaludwigphoto  
<https://www.flickr.com/photos/saschaludwig/20194395853>

**EXIT**  
**HAVE A SAFE**  
**JOURNEY HOME**

Image taken in on HMB Endeavour (replica) in Sydney/NSW

[https://fosdem.org/2018/schedule/event/community\\_passing\\_the\\_button\\_foss\\_leadership/](https://fosdem.org/2018/schedule/event/community_passing_the_button_foss_leadership/)

People

Licensing

Trademarks

Patents

Marketing

Education

Documentation

Design

Event management

Social media

Support

Funding

Motivation



<http://foss-backstage.de>  
Berlin, 2019



<http://www.apachecon.com>  
Montreal, Sep 2018

Join:

`dev@your-favourite-project.apache.org`

`user@your-favourite-project.apache.org`

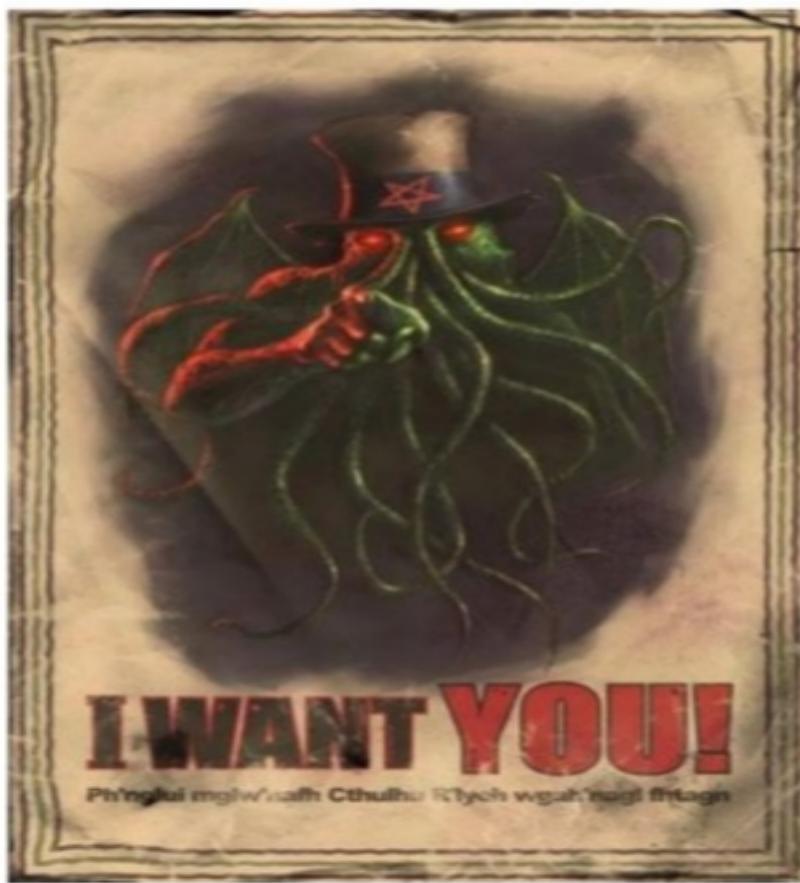


Image by: Patrick McEvoy

Bug reports, patches, features.  
Documentation, code, examples.

<http://s.apache.org/e-p>



Join:

`dev@your-favourite-project.apache.org`

`user@your-favourite-project.apache.org`



Image by: Patrick McEvoy

Bug reports, patches, features.  
Documentation, code, examples.