

# Webkit Project – A FLOSS report

## Dynamics of Libre Software Communities

Simón Pena Placer

Máster Software Libre, 2009-2010. A Coruña Edition

January 21, 2010

# License



This work is licensed under the Creative Commons Attribution-Share Alike 3.0 Spain License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/3.0/es/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

# Content

- 1 Introduction
- 2 Methodology and tools
- 3 Results
  - Source code repository
  - Developers mailing list
  - Unassigned bugs mailing list
- 4 Conclusions

# Introduction (I)

## About this work

- An Open Source Project will be analyzed
- Which one? *Why?*
- WebKit project was chosen
- There is a technology war between web browsers
  - Proprietary browser Internet Explorer rules the market
  - However, *times may be changing*: Open Source browsers share is growing

# Introduction (I)

## About this work

- An Open Source Project will be analyzed
- Which one? Why?
- WebKit project was chosen
- There is a technology war between web browsers
  - Proprietary browser Internet Explorer rules the market
  - However, *times may be changing*: Open Source browsers share is growing

# Introduction (I)

## About this work

- An Open Source Project will be analyzed
- Which one? Why?
- WebKit project was chosen
- There is a technology war between web browsers
  - Proprietary browser Internet Explorer rules the market
  - However, *times may be changing*: Open Source browsers share is growing

# Introduction (I)

## About this work

- An Open Source Project will be analyzed
- Which one? Why?
- WebKit project was chosen
- There is a technology war between web browsers
  - Proprietary browser Internet Explorer rules the market
  - However, *times may be changing*: Open Source browsers share is growing

# Introduction (I)

## About this work

- An Open Source Project will be analyzed
- Which one? Why?
- WebKit project was chosen
- There is a technology war between web browsers
  - Proprietary browser Internet Explorer rules the market
  - However, *times may be changing*: Open Source browsers share is growing



# Introduction (I)

## About this work

- An Open Source Project will be analyzed
- Which one? Why?
- WebKit project was chosen
- There is a technology war between web browsers
  - Proprietary browser Internet Explorer rules the market
  - However, *times may be changing*: Open Source browsers share is growing

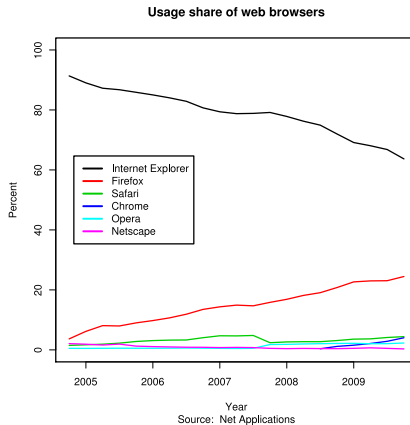
# Introduction (I)

## About this work

- An Open Source Project will be analyzed
- Which one? Why?
- WebKit project was chosen
- There is a technology war between web browsers
  - Proprietary browser Internet Explorer rules the market
  - However, *times may be changing*: Open Source browsers share is growing

# Introduction (II)

## The Browser War



# Introduction (II)

## Beyond the browser war

- The war has been extended to applications using web browser's engines.
- Gecko, Mozilla's engine, is one of the most used.
- WebKit, Apple's Safari engine, is experimenting an enormous growth

# Introduction (III)

## About this work

- *Which engine is better suited to your application development?*
- This work will help you decide, by analyzing the WebKit project.
- A compared analysis with Gecko is out of the scope of this work
  - Gecko's code is too coupled with Mozilla's
  - It is hard to analyze it alone

# Introduction (IV)

Some info about the WebKit project

- WebKit is Apple's Safari layout engine
- It started from KDE's KHTML and KJS engines
- Its development has been opening more and more
- Nowadays its code is under the LGPL or Apache Licenses

# Introduction (and V)

## Projects using WebKit

There are many projects using WebKit. Regarding to Web browsers:

- Safari and Safari Mobile, from Apple's OS X and iPhone OS, respectively
- Google Chrome Browser and Chromium Open Source project
- Android's Web browser
- Nokia's S60 browser
- Epiphany, who has moved from Gecko
- Others. . .

# Methodology

The following project resources will be analyzed

- Source code repository, located at <http://svn.webkit.org/repository/webkit/trunk/>
- Developers mailing list, located at <https://lists.webkit.org/pipermail/webkit-dev/>
- Unassigned bugs mailing list, located at <https://lists.webkit.org/pipermail/webkit-unassigned/>



# Tools

- Libresoft tools[1] will be used
  - cvsanaly: <http://git.libresoft.es/cvsanaly/>
  - mlstats:  
<http://forge.morfeo-project.org/plugins/scmsvn/viewcvs.php/maillingliststat/?root=libresoft-tools>
  - guilty: <http://git.libresoft.es/guilty/>
- Project data will be exported into a database via these tools
- A series of queries will be used in order to gather relevant information
- R[2] will be used to issue the queries and export their results
- Some queries from Flossreport were also used

# Source Code Repository

- Most important info will be summarized: number of commits and committers, lines of code, etc.
- Most active committers will be identified, both all-time and last year
- Companies behind development will be identified, and its relevance evaluated
- Information about trends will be extracted, different periods will be shown

# Developers Mailing List

- Most important info will be summarized: number of emails sent and writers, lifespan, etc.
- Most active posters will be identified, both all-time and last year
- Companies behind development will be identified, and its relevance evaluated
- Information about trends will be extracted, different periods will be shown.

# Unassigned bugs mailing list

- All-time number of emails received will be shown
- Mailing list trends will be exposed

# Outline

- 1 Introduction
- 2 Methodology and tools
- 3 **Results**
  - Source code repository
  - Developers mailing list
  - Unassigned bugs mailing list
- 4 Conclusions

# Summary

| Concept                               | Count  |
|---------------------------------------|--------|
| Number of commits                     | 44143  |
| Number of committers                  | 195    |
| Number of files under version control | 80860  |
| Number of lines                       | 824955 |
| Years of activity                     | 9      |

**Table:** Brief summary of the repository's activity

# Programming language distribution

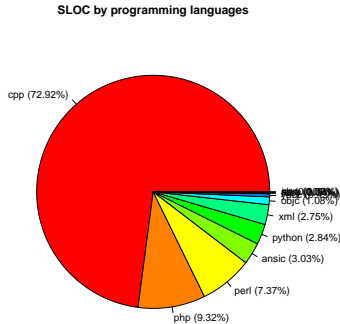


Figure: Programming language distribution

# All-time top 10 committers

|    | Committer           | Commit count |
|----|---------------------|--------------|
| 1  | darin               | 3583         |
| 2  | hyatt               | 2158         |
| 3  | eric@webkit.org     | 1967         |
| 4  | mjs                 | 1620         |
| 5  | hausmann@webkit.org | 1174         |
| 6  | rjw                 | 1104         |
| 7  | darin@apple.com     | 1076         |
| 8  | kocienda            | 958          |
| 9  | mitz@apple.com      | 945          |
| 10 | mrowe@apple.com     | 941          |

**Table:** Top 10 committers. Multiple accounts ignored



## All-time top 10 committers – accounts grouped

|    | Committer | Commit count |
|----|-----------|--------------|
| 1  | darin     | 4876         |
| 2  | hyatt     | 3042         |
| 3  | eric      | 1967         |
| 4  | mjs       | 1842         |
| 5  | hausmann  | 1417         |
| 6  | andersca  | 1376         |
| 7  | ggaren    | 1287         |
| 8  | weinig    | 1263         |
| 9  | ap        | 1252         |
| 10 | aroben    | 1164         |

**Table:** Top 10 committers. Multiple accounts grouped

## 2009 top 10 committers

|    | Committer              | Commit count |
|----|------------------------|--------------|
| 1  | eric@webkit.org        | 1646         |
| 2  | abarth@webkit.org      | 474          |
| 3  | hausmann@webkit.org    | 440          |
| 4  | kov@webkit.org         | 387          |
| 5  | darin@apple.com        | 385          |
| 6  | mrowe@apple.com        | 364          |
| 7  | simon.fraser@apple.com | 357          |
| 8  | mitz@apple.com         | 334          |
| 9  | hyatt@apple.com        | 322          |
| 10 | oliver@apple.com       | 320          |

**Table:** Top 10 committers during 2009

# Commits by company

|   | Company         | First contribution | Last contribution | Commit count |
|---|-----------------|--------------------|-------------------|--------------|
| 1 | apple.com       | 2007               | 2010              | 12005        |
| 2 | webkit.org      | 2007               | 2010              | 8753         |
| 3 | chromium.org    | 2008               | 2009              | 1831         |
| 4 | nokia.com       | 2009               | 2010              | 63           |
| 5 | google.com      | 2009               | 2009              | 30           |
| 6 | torchmobile.com | 2009               | 2009              | 10           |
| 7 | forwardbias.in  | 2009               | 2009              | 8            |

**Table:** Number of commits by the company an user is affiliated to

# Yearly activity

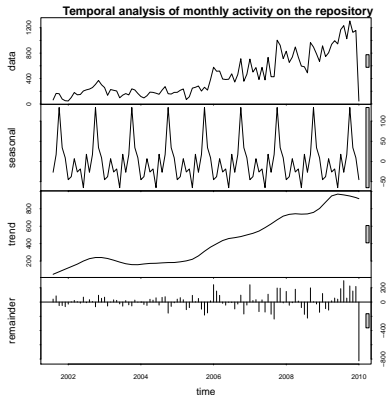


Figure: Evolution of yearly activity in the repository

# Weekly activity

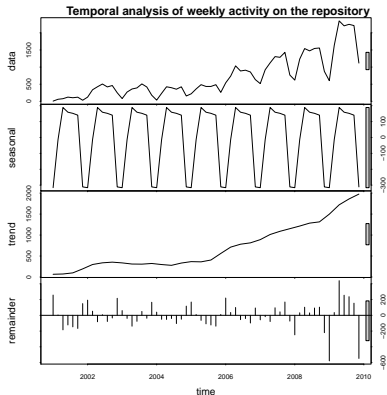


Figure: Evolution of weekly activity in the repository

# Daily activity

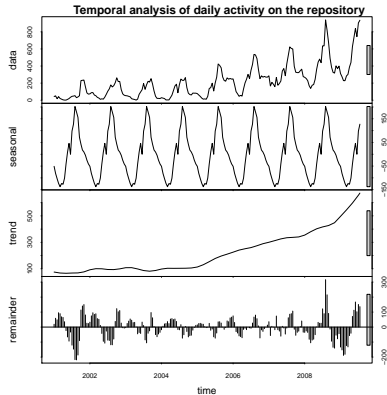
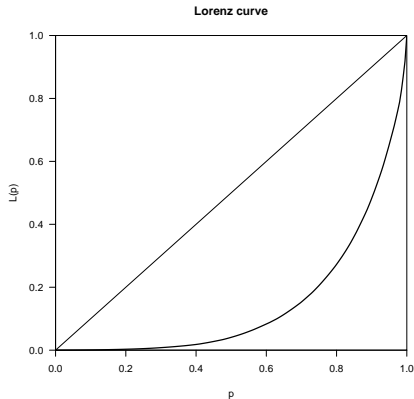


Figure: Evolution of daily activity in the repository

# Gini coefficient and Lorenz curve



**Figure:** Lorenz curve for the repository. Gini coefficient is 0.706

# Outline

- 1 Introduction
- 2 Methodology and tools
- 3 Results**
  - Source code repository
  - Developers mailing list**
  - Unassigned bugs mailing list
- 4 Conclusions



# Summary

| Concept                                    | Count |
|--|-------|
| Emails sent to the list                    | 9760  |
| Unique email addresses writing to the list | 1199  |
| Different user names writing to the list   | 1132  |
| Years of activity                          | 5     |

[Table:](#) Brief summary of the developers' mailing list

## All-time top 10 posters

|    | Username   | Email count |
|----|------------|-------------|
| 1  | darin      | 623         |
| 2  | mjs        | 605         |
| 3  | ddkilzer   | 307         |
| 4  | mrowe      | 209         |
| 5  | aroben     | 206         |
| 6  | mike.emmel | 173         |
| 7  | eric       | 170         |
| 8  | hyatt      | 165         |
| 9  | ggaren     | 160         |
| 10 | abarth     | 147         |

Table: Top 20 posters

## 2009 top 10 posters

|    | Username | Email count |
|----|----------|-------------|
| 1  | darin    | 288         |
| 2  | mjs      | 233         |
| 3  | abarth   | 132         |
| 4  | eric     | 123         |
| 5  | pkasting | 106         |
| 6  | jorlow   | 96          |
| 7  | ddkilzer | 86          |
| 8  | mrowe    | 78          |
| 9  | ggaren   | 72          |
| 10 | aroben   | 69          |

**Table:** Top 20 posters during 2009

# Messages by company

|    | Domain name  | Email count |
|----|--------------|-------------|
| 1  | apple.com    | 2329        |
| 2  | gmail.com    | 2244        |
| 3  | webkit.org   | 665         |
| 4  | chromium.org | 423         |
| 5  | google.com   | 337         |
| 6  | yahoo.com    | 221         |
| 7  | mac.com      | 169         |
| 8  | kde.org      | 134         |
| 9  | kilzer.net   | 110         |
| 10 | selfish.org  | 91          |

**Table:** Number of emails sent by company employees

# Yearly activity

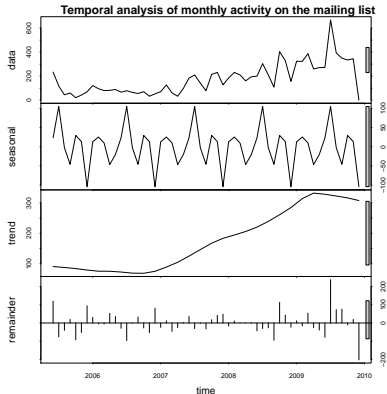


Figure: Evolution of yearly activity in the developers' mailing list

# Weekly activity

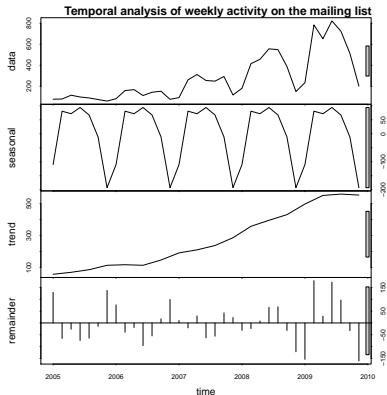


Figure: Evolution of weekly activity in the developers' mailing list

# Daily activity

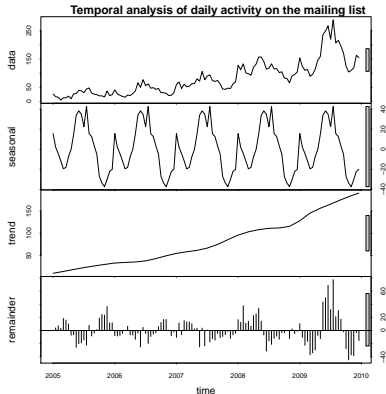


Figure: Evolution of daily activity in the developers' mailing list

# Outline

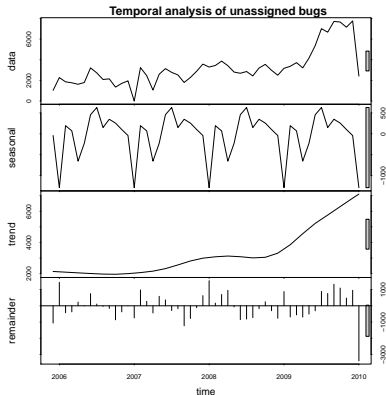
- 1 Introduction
- 2 Methodology and tools
- 3 Results**
  - Source code repository
  - Developers mailing list
  - **Unassigned bugs mailing list**
- 4 Conclusions



# Summary

- Each time a bug is created, a message is sent to the list
- Each time a bug changes its state, another message is sent
- 160342 Emails sent to the list

# Yearly activity



**Figure:** Evolution of the monthly activity on the unassigned bugs mailing list

# Conclusions

Some work has been done...

An analysis about WebKit project has been performed:

- The overall trend for the project's growth is clearly positive
- Most important contributors have been identified for both the repository and mailing list
- The development model has been exposed
- Companies supporting the development have been identified

# Conclusions (and II)

...but there is still work left to do

- Handling WebKit project as a whole makes it harder to understand
  - Analysis performed on smaller parts would be interesting
- Different queries could show different information currently omitted
- Other layouts should be taken into account for the analysis to be completed
  - Specific tests should be designed in order to help comparing engines
  - Gecko engine would be the first candidate



## LibreSoft.

start [GSyC/LibreSoft tools].

<http://tools.libresoft.es/>, 2010.



## R Development Core Team.

*R: A Language and Environment for Statistical Computing.*

R Foundation for Statistical Computing, Vienna, Austria,  
2009.

ISBN 3-900051-07-0.



## WebKit.

The WebKit Open Source Project.

<http://webkit.org>, 2010.

# Webkit Project – A FLOSS report

## Dynamics of Libre Software Communities

Simón Pena Placer

Máster Software Libre, 2009-2010. A Coruña Edition

January 21, 2010