

**SIXTH FRAMEWORK PROGRAMME**  
**PRIORITY 2**  
**INFORMATION SOCIETY TECHNOLOGIES**



**SIXTH FRAMEWORK  
PROGRAMME**

**FLOSSWORLD**

**Free/Libre and Open Source Software: Worldwide  
Impact Study**



Track 2 Study Report – South Africa  
(Referred to as D29b in the work packages description in the proposal)

Project Reference: 015722

Kind of Project: Specific Support Action

Start Date: May 2005

End Date: June 2007



UNITED NATIONS  
UNIVERSITY

**UNU-MERIT**



Universidad  
Rey Juan Carlos

## **Authors and contributors**

### *Coordination, Drafting and Edition*

Jesús M. González-Barahona, GSyC/Libresoft, Universidad Rey Juan Carlos

Daniel Izquierdo-Cortázar, GSyC/Libresoft, Universidad Rey Juan Carlos

Álvaro del Castillo, GSyC/Libresoft, Universidad Rey Juan Carlos

### *Reviewing*

Gregorio Robles, GSyC/Libresoft, Universidad Rey Juan Carlos

Teófilo Romera, GSyC/Libresoft, Universidad Rey Juan Carlos

### *Retrieval process and Drafting*

Roberto Andradás-Izquierdo, GSyC/Libresoft, Universidad Rey Juan Carlos

Jorge Gascón-Pérez, GSyC/Libresoft, Universidad Rey Juan Carlos

Daniel Izquierdo-Cortázar, GSyC/Libresoft, Universidad Rey Juan Carlos

Juan Luis Prieto, GSyC/Libresoft, Universidad Rey Juan Carlos

### *Graphics*

Juan Martínez-Romo, GSyC/Libresoft, Universidad Rey Juan Carlos

### *Local support and Reviewing*

Pumeza Ceza, The Council for Scientific and Industrial Research

Derek Keats, The University of the Western Cape

Juliet Lesch, The University of the Western Cape

Nhlanhla Mabaso, The Council for Scientific and Industrial Research

Kim Tucker, The Council for Scientific and Industrial Research

Enver Ravat, The University of the Western Cape

## **Copyright**

©2007 GSyC/LibreSoft

Some rights reserved. This report is distributed under the Creative Commons Attribution-ShareAlike 3.0 licence, available in <http://creativecommons.org/licenses/by-sa/3.0>

This report is available in <http://flossworld.org>

This report has been funded by the European Commission under contract number FP6-IST-015722.

## **Disclaimer**

The opinions expressed in this Study are those of the authors and do not necessarily reflect the views of the European Commission. Contract FP6-IST-015722

# Contents

<b>1</b>	<b>Executive Summary</b>	<b>5</b>
<b>2</b>	<b>Introduction</b>	<b>6</b>
2.1	Details of the considered region . . . . .	6
2.2	Summary of results . . . . .	7
2.3	Problems and constraints found . . . . .	9
2.4	Methodology details . . . . .	10
2.5	Contributors . . . . .	10
<b>3</b>	<b>Results</b>	<b>12</b>
3.1	General information . . . . .	12
3.2	Forges information . . . . .	13
3.3	SCM information . . . . .	13
3.3.1	Avoir Forge . . . . .	13
3.3.2	Knowledge Tree forge . . . . .	15
3.4	Source code information . . . . .	15
3.4.1	Avoir forge . . . . .	15
3.4.2	Knowledge Tree forge . . . . .	17
3.4.3	Aggregated data . . . . .	19
3.5	Mailing Lists . . . . .	20
3.6	Authorship information . . . . .	20
3.7	Sourceforge Analysis . . . . .	21
<b>4</b>	<b>Appendixes</b>	<b>23</b>
4.1	General information . . . . .	23
4.2	Data from repositories . . . . .	25
4.3	COCOMO model applied to software releases . . . . .	31

# List of Tables

2.1	South African forges (on 16th of April of 2007)	7
2.2	South African forges (on April-June of 2007). SourceForge has been included for completeness (data June, 2006)	8
3.1	Data provided by South African partners	12
3.2	South African forges	13
3.3	Avoir committers and commits	14
3.4	Projects with the most active developers	14
3.5	Committers with the highest number of commits	14
3.6	Project vs Number of committers and number of commits	15
3.7	Basic COCOMO model applied (using SLOCCount) to the largest Avoir forge software releases	15
3.8	Programming languages used in Avoir forge	16
3.9	Basic COCOMO model applied (using SLOCCount) to the largest Knowledge Tree forge software releases	17
3.10	Programming languages used in Knowledge Tree forge	18
3.11	Projects with higher effort	19
3.12	Programming languages used in South African forges	20
3.13	General mailing lists summary of results	20
3.14	Authorship results	21
3.15	General results from Sourceforge	21
3.16	The most active South African committers detected in Sourceforge	22
4.1	Communities detected by partners	23
4.2	South African LUGs	23
4.3	Projects collected by partners	24
4.4	Platforms collected by partners	24
4.4	Platforms collected by partners	25
4.5	SCM results for Avoir Forge	25
4.5	SCM results for Avoir Forge	26
4.5	SCM results for Avoir Forge	27
4.5	SCM results for Avoir Forge	28
4.6	SCM results in SourceForge forge	29
4.6	SCM results in SourceForge forge	30
4.6	SCM results in SourceForge forge	31
4.7	Basic COCOMO model applied to Avoir software distributions	32
4.8	Basic COCOMO model applied to Knowledge Tree forge software distributions	32
4.8	Basic COCOMO model applied to Knowledge Tree forge software distributions	33

# List of Figures

2.1	South Africa World Map . . . . .	7
2.2	SCM repositories, mailing lists and software releases found in forges . . . . .	8
2.3	Committers per forge . . . . .	9
2.4	Commits per forge . . . . .	9
3.1	General information obtained from surveys (comparison among countries) . . . . .	13
3.2	Distribution of commits per project in Avoir forge . . . . .	14
3.3	Lines of code in Avoir forge . . . . .	16
3.4	Programming languages used in Avoir forge . . . . .	17
3.5	Lines of code in Knowledge Tree forge . . . . .	18
3.6	Programming languages used in Knowledge Tree forge . . . . .	19
3.7	Authorship data (in software releases). . . . .	21
3.8	Commits in SourceForge . . . . .	22

# Chapter 1

## Executive Summary

This report shows the main results of the analysis performed, in the context of the FLOSSWorld project, of some quantitative aspects of libre (free, open source) software in South Africa. The sources of the analysis have been an exhaustive data retrieval of several facts related to libre software in that country (such as a list of Linux user groups, magazines focused on libre software, etc.) and a detailed, quantitative data mining of various sites hosting libre software development (forges).

In South Africa we can find 2 local forges (Avoir and Knowledge Tree). These two forges host the development of libre software projects in South Africa. There is also a community of people identified as South Africans who use SourceForge for the same purpose, and this fact has been taken into account in this report.

Regarding the two local forges, Knowledge Tree appears to be the one with the most registered users, around 310. On the other hand, we could not find out how many registered users Avoir has. However Avoir is the forge with the most registered projects. Avoir has 70 registered projects and Knowledge Tree has 51. The data from South African activity in SourceForge shows around 5700 South African users registered there and nearly 500 projects which were identified as South African-driven.

The data gathered during Source Code Management systems mining, show that Avoir has 22 repositories and 71 different committers using them. In this case, Knowledge Tree is a black box because automatic data retrieval is not possible in their Source Code Management systems. Moreover, just 73 of the 494 South African projects in SourceForge use the Source Code Management system tool provided by the platform. Despite the difference of the number of repositories mined in Avoir (22) and in SourceForge (73), there is much more source code activity in Avoir than in SourceForge's South African community. Avoir has a history of around 92,888 commits while SourceForge shows 53,142.

It is also remarkable that the most popular programming language in the South African local forges is, by far, php. Php represents 82% of the source code in Avoir (followed by python) and 98% in Knowledge Tree (followed by C#). The total lines of code the two forges sum is 244,222.

Some other relevant facts include that authorship data shows 52 different individual authors, 4 universities and 9 enterprises. Or the sterile creation of the mailing lists in Avoir (14 in total) that have never been used.

# Chapter 2

## Introduction

Within the context of the FLOSSWorld project, this report is devoted to the quantitative study of libre (free, open source) software development in South Africa. The information presented here is based on public data found in the repositories of libre software projects hosted in the analysed area, in SourceForge (the largest hosting site for libre software projects) and in a survey completed with the help of the partners in FLOSSWorld.

The data found in repositories (usually identified by the local partners in the project) has been downloaded, stored in a database, and later carefully mined and analysed (using a semi-automatic process that has been complemented by human validation). Most of the data used was obtained from source code, source code management systems and mailing list archives. In addition to repositories in the region, SourceForge has also been analysed, as the largest hosting site, worldwide, for libre software projects. The details of the methodology used are specified in a separate document, also produced by the FLOSSWorld project, the “Methodology report”.

This document shows the main results produced by this methodology. Before that, this chapter discusses some details about the methodology itself.

### 2.1 Details of the considered region

**Geographic area.** The Republic of South Africa<sup>1</sup> is located in the south of African continent and its capital is divided in three cities. Pretoria is the executive capital, Bloemfontein is the judicial capital and Cape Town is the legislative capital, being this one the official capital. Its population is estimated in 47.432.000 people and its estimated GDP is \$570,2 billion (\$12,161 per capita).

**Languages.** South Africa has 11 official languages, Afrikaans, English, Ndebele, Northern Sotho, Sotho, Swait, Tsonga, Tswana, Venda, Xhosa and Zulu<sup>2</sup>.

**Map.** Figure 2.1 shows the location of South Africa.

---

<sup>1</sup>[http://en.wikipedia.org/wiki/South\\_Africa](http://en.wikipedia.org/wiki/South_Africa)

<sup>2</sup>information based on Wikipedia [http://en.wikipedia.org/wiki/Languages\\_of\\_South\\_Africa](http://en.wikipedia.org/wiki/Languages_of_South_Africa)



Figure 2.1: South Africa World Map

**Time zone.** The South African time zone<sup>3</sup> is SAST (UTC +2) and it is relevant for this study; SAST is a time zone used by all of South Africa, in addition to Lesotho and Swaziland. The zone is two hours ahead of UTC (UTC+2), which is the same as Central Africa Time, used in countries to the north and east of South Africa.

**Internet top level domain.** South African TLD<sup>4</sup> (Internet Top Level Domain which is *.za*) is an important indicator for information gathering. When recovering information from email addresses, if a developer or user has a South African TLD, we can reasonably assume that she is from India.

[Both the description of the geographic area and the map of South Africa where obtained in the Wikipedia<sup>5</sup>]

## 2.2 Summary of results

The South African FLOSSWorld partners have identified two forges which should be considered as South African local forges. These forges have been spidered in order to identify the projects they contain. Table 2.1 lists a relation of the number of projects and users registered in each of the forges. The world's most popular forge, SourceForge, has been added to the table as many South African developers and South African-driven projects have been found there. The number of registered users in SourceForge should be understood as the estimation of South African developers<sup>6</sup> in SourceForge. The 494 projects in SourceForge that are identified as South African-driven have a majority (i.e. more than 50%) of South African developers in their teams.

Forge	Registered Users	Registered Projects
Avoir	no data	70
Knowledge Tree	308	51
Sourceforge	5706	494

Table 2.1: South African forges (on 16th of April of 2007)

It is important to point out that not all registered users are active developers in the forges. Many of them could register and never join a development project, for instance. Projects, as well, may not make use of all development-related tools offered by the forges (and therefore, some kinds of data sources, may offer no data at all). This fact is clearly revealed in the summary table 2.2. This table ( 2.2) shows the number of SCM (CVS/SVN) repositories, committers, commits, mailing lists, software releases and size of the software.

<sup>3</sup>Methodology report - Chapter Methodology - Section Global forge's analysis

<sup>4</sup>Methodology report - Chapter Methodology - Section Global forge's analysis

<sup>5</sup>[http://en.wikipedia.org/wiki/South\\_Africa](http://en.wikipedia.org/wiki/South_Africa)

<sup>6</sup>Methodology report: How to obtain nationality from Sourceforge

Forge	SCM Repos.	Committers	Commits	MailingLists	Releases	SLOC
Avoir	22	71	92,888	14	8	203,286
Knowledge Tree	ND	ND	ND	ND	39	40,936
Sourceforge	73	102	53,142	ND	ND	ND

Table 2.2: South African forges (on April-June of 2007). SourceForge has been included for completeness (data June, 2006)

The next section describes the set of difficulties arisen during the data retrieval process. These difficulties explain the *ND* (No-Data) value that appears in some cells of the table 2.2.

In the results shown in table 2.2 the field *SCM repositories* (Source Code Management repositories) presents the number of non empty repositories and the number of total repositories found (including both empty and non empty repositories<sup>7</sup>). Field *Releases* shows the number of software releases which could be analysed and the number of total software releases found. For SourceForge the result presented is the number of non empty SCM repositories. Finally, field *SLOC* means “Source Lines Of Code”.

Figure 2.2 shows the number of SCM repositories, mailing lists and software releases identified and analysed.

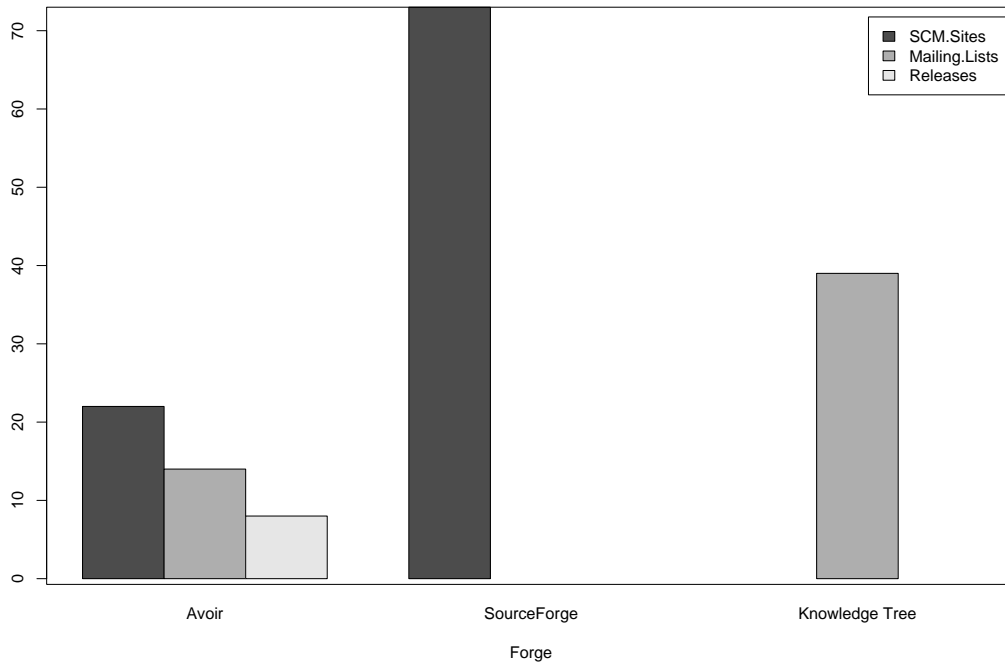


Figure 2.2: SCM repositories, mailing lists and software releases found in forges

<sup>7</sup>Empty repository: There is a valid SCM account for this project, but no commit has been made

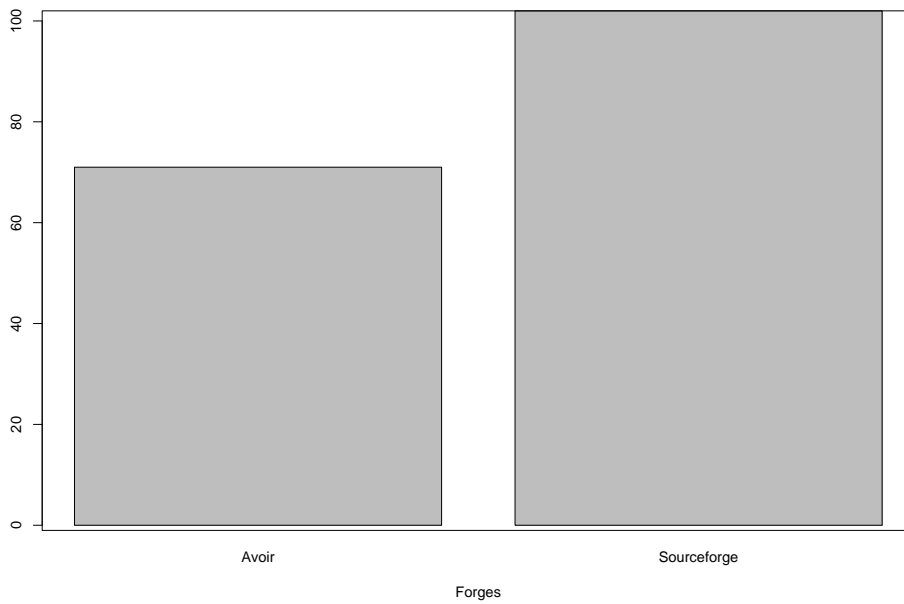


Figure 2.3: Committers per forge

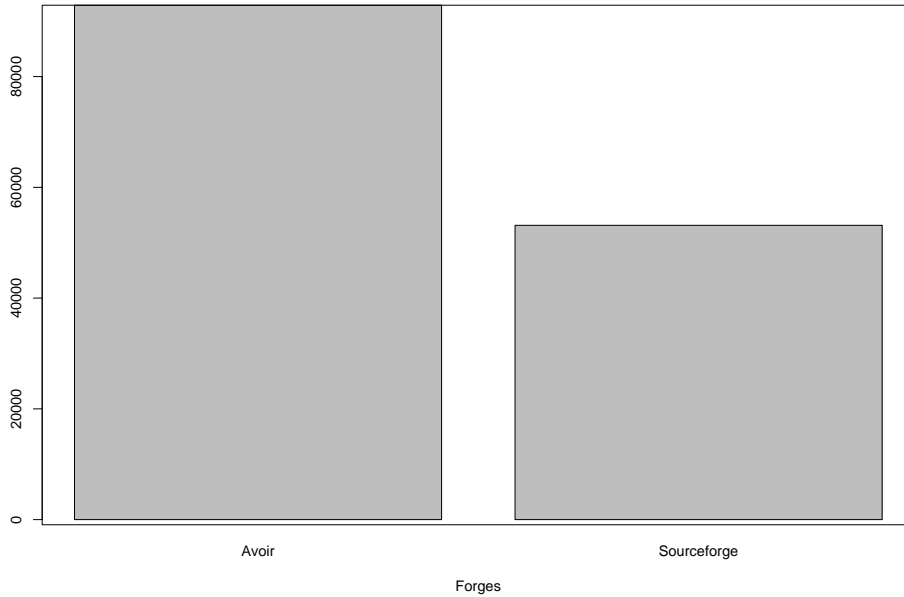


Figure 2.4: Commits per forge

## 2.3 Problems and constraints found

The data retrieval process presented two main problems:

1. Links to version control systems: South African partners provided some useful links. But sometimes it was not enough to find repositories (CVS or SVN) links, thus we spent some time looking for specific links we needed. At the end we were able to find extra information in forges using our spider tool.
2. Empty software repositories: Some projects which did have a CVS or SVN repository, did not store any source code in them. In this cases, some manual research and tests helped to clarify that repositories where actually empty.

In addition to the aforementioned general problems, there have been specific issues on some of the forges which are explained next:

1. Avoir forge: The Spider tool could not find any SMC sites. However, the partners provided the SCM repository URL and data could be retrieved then.
2. Knowledge Tree forge: This forge needs an account in order to download repositories from its server. Its web page says that there is an anonymous access, however it was not possible to download any repository. After sending an email to the administrators of Knowledge Tree forge, the problem is not yet fixed.

## 2.4 Methodology details

This section presents a brief description of the different tools used in the data retrieval process. Three forges have been analysed, Sourceforge and Avoir. This is why data from several sources is available. The analysis was carried out on January of 2007. The study involved different tools in order to automate the gathering of data. Sourceforge analysis was run in June 2006.

1. CVSanaly<sup>8</sup>: Knowledge Tree forge could not be analysed.
2. MailingListStats<sup>9</sup>: It was run on mailing lists from local forges but most of them were empty. Empty mailing lists are those that were created once, but there are not messages in their history.
3. Pyternity<sup>10</sup>: It was run on local forge's software releases (not on SCMs).

Knowledge Tree is a document management system. It has its own forge which collects all projects related to Knowledge tree. This forge (with its projects) is located in South Africa so it will be considered as a South African forge, even language package projects. However repositories from this forge need authentication, so it was not possible to analyze this forge in depth. At the end there are some general results about number of mailing lists, number of releases or number of repositories.

## 2.5 Contributors

This report has been drafted by the GSyC/LibreSoft team of the Universidad Rey Juan Carlos (Madrid, Spain, coordinated by Jesus M. Gonzalez-Barahona). The URJC team performed the data retrieval, mining and analysis, and produced its final version. In addition to the general collaboration by all partners, two South African FLOSSWorld partners have contributed specifically (The Council for Scientific

---

<sup>8</sup>Methodology report - Chapter Tools - Section CVSanaly

<sup>9</sup>Methodology report - Chapter Tools - Section Mailing List Stats

<sup>10</sup>Methodology report - Chapter Tools - Section Pyternity

and Industrial Research -<http://www.csir.co.za/>- and The University of the Western Cape - <http://www.uwc.ac.za/>-) by providing information about libre software developing sites, feedback about the drafts of the report, much other useful information, and generic help with issues specific to their region. The coordinator of FLOSSWorld (UNU-MERIT) has also provided specific assistance and feedback during all the stages of preparation of this report.

## Chapter 3

# Results

### 3.1 General information

South Africa partners collaborated with us providing different sources of useful data for the report<sup>1</sup>. For more information access section 4.1 where this data is detached.

There is a summary in table 3.1.

<b>N Communities</b>	<b>N Lugs</b>	<b>Platforms</b>	<b>Projects</b>
3	3	14	21

Table 3.1: Data provided by South African partners

Where *Communities* means group of users interested in libre software, *LUGs* are Oficial Linux User Groups (with a physical address) and *Platforms* are web sites which provide any kind of support to the libre software world, such as forges. Also, in figure 3.1 there is a comparison of the information obtained from all the countries studied in FLOSSWorld.

---

<sup>1</sup>Methodology report - Chapter Data Sources - Section Primary Data Sources

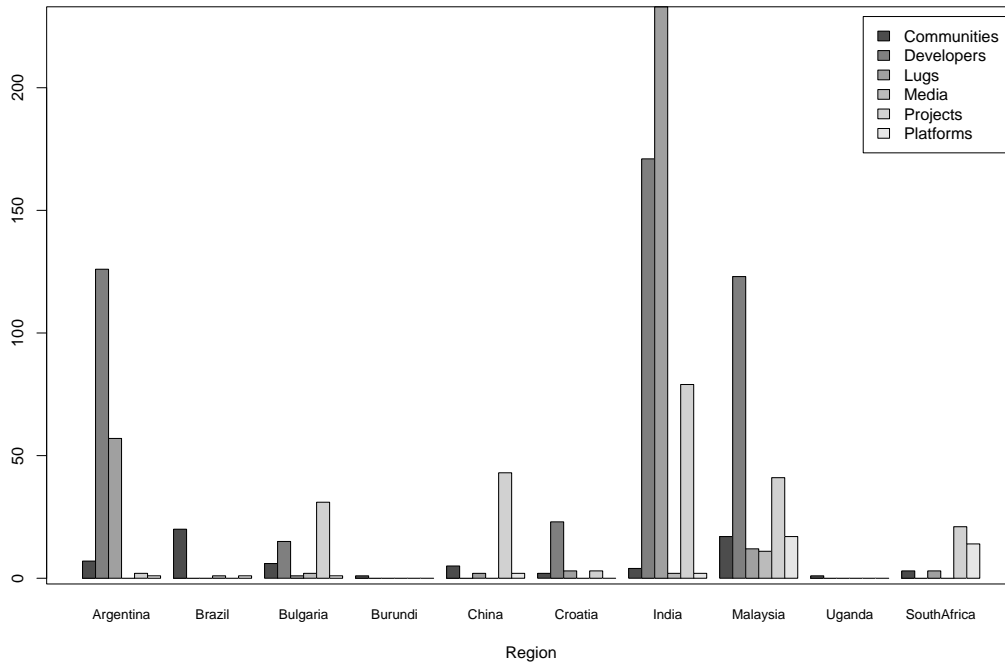


Figure 3.1: General information obtained from surveys (comparison among countries)

## 3.2 Forges information

Two forges have been found. Avoir forge<sup>2</sup> and Knowledge Tree forge<sup>3</sup>. The first one is a general purpose forge created to be a local forge in South Africa. The second forge was created in order to store projects related to the Knowledge Tree project<sup>4</sup>

Forge	Forge name
<a href="http://gforge2.uwc.ac.za/">http://gforge2.uwc.ac.za/</a>	Avoir
<a href="http://www.knowledgetree.com">http://www.knowledgetree.com</a>	Knowledge Tree

Table 3.2: South African forges

## 3.3 SCM information

### 3.3.1 Avoir Forge

70 projects have been found, but only 22 of them have data. There are 71 active developers who have made at least a commit and in total we can account up to 92,888 commits (see table 3.3).

<sup>2</sup><http://gforge2.uwc.ac.za/>

<sup>3</sup><http://forge.ktdms.com/>

<sup>4</sup><http://www.knowledgetree.com/>

Forge	No. of Committers	No. of Commits
Avoir	71	92,888

Table 3.3: Avoir committers and commits

Table 3.4 shows the five most active committers in projects in Avoir forge. Also, table 3.5 shows the most active developers in Avoir forge (with most of the commits).

Project	Committers	Commits
nextgen	tohir	9679
nextgen	dkeats	7186
odimex	dkeats	6712
nextgen	joconnor	5060
nextgen	megan	4333

Table 3.4: Projects with the most active developers

Committers	Commits
tohir	14433
dkeats	14078
joconnor	8034
wnitsckie	6650
megan	6406

Table 3.5: Committers with the highest number of commits

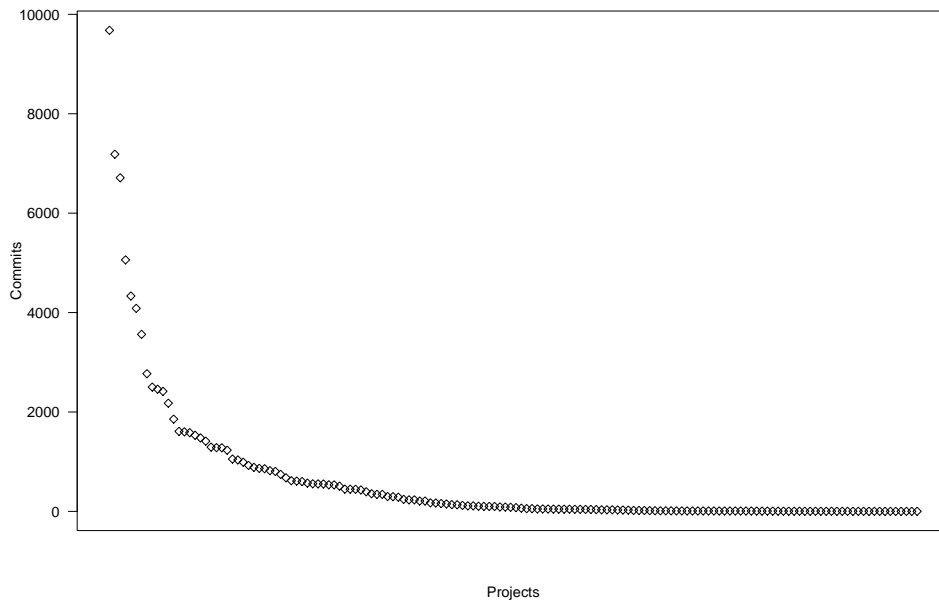


Figure 3.2: Distribution of commits per project in Avoir forge

Project Name	Number of Committers	Number of Commits
nextgen	56	51706
odimex	25	22678
sems	10	7180
chisimba_skins	20	5399
zinf	3	1832

Table 3.6: Project vs Number of committers and number of commits

### 3.3.2 Knowledge Tree forge

No data could be retrieved from this one. It is mandatory to have an account (as developer) for each project. So the one-by-one approach is not possible. And the anonymous access was not work properly, even after contacting the administration team of the forge.

## 3.4 Source code information

The data for the analysis of the source code was obtained by downloading releases of software and then running SLOCCount on them. This whole process was automatic and carried out by our spider software.

### 3.4.1 Avoir forge

Table 3.7 and figure 3.3 show results from applying the basic COCOMO model for cost estimation, to releases obtained from Avoir.

Project name	No. of SLOC	Effort	Schedule	No. developers	Estimated cost
nextgen	164,634	42.50 (509.98)	2.23 (26.72)	19.09	\$ 5,740,994
exe	15,304	3.51 (42.10)	0.86 (10.36)	4.07	\$ 473,900
zinf	13,873	3.16 (37.97)	0.83 (9.96)	3.81	\$ 427,485
c2go	3,243	0.69 (8.25)	0.46 (5.58)	1.48	\$ 92,926
ziggis	3,101	0.66 (7.88)	0.46 (5.48)	1.44	\$ 88,658

Table 3.7: Basic COCOMO model applied (using SLOCCount) to the largest Avoir forge software releases

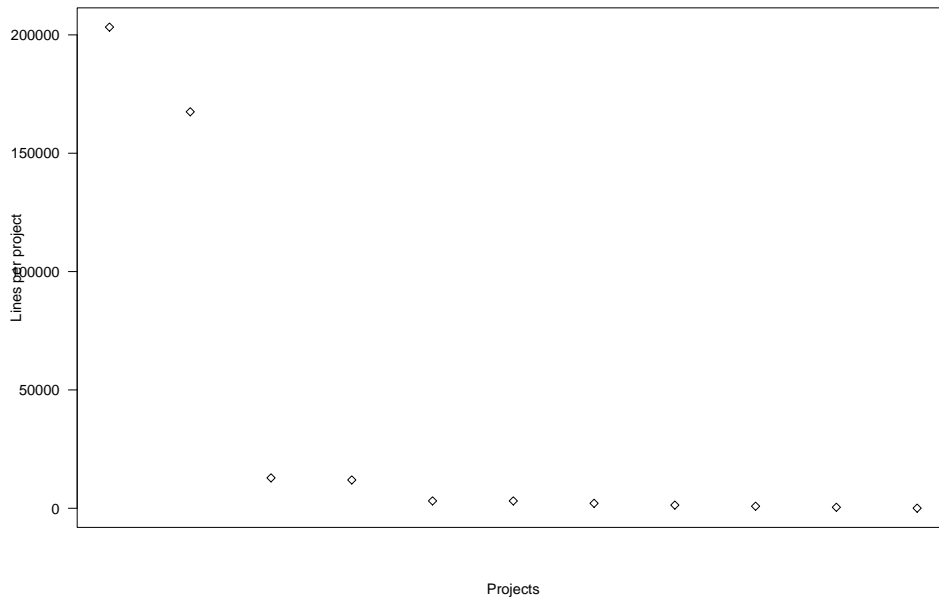


Figure 3.3: Lines of code in Avoir forge

Also, table 3.8 and figure 3.4 show information regarding to percentage of use of different programming languages in the Avoir forge.

Programming language	No. of lines	Percentage
PHP	167512	82.4021
Python	12837	6.3147
Java	11964	5.8853
C	3113	1.5313
C #	3101	1.5254
JSP	2092	1.0290
SH	1337	0.6576
Perl	862	0.4240
Pascal	44	0.0216
C ++	424	0.2085
Total	203286	100.0

Table 3.8: Programming languages used in Avoir forge

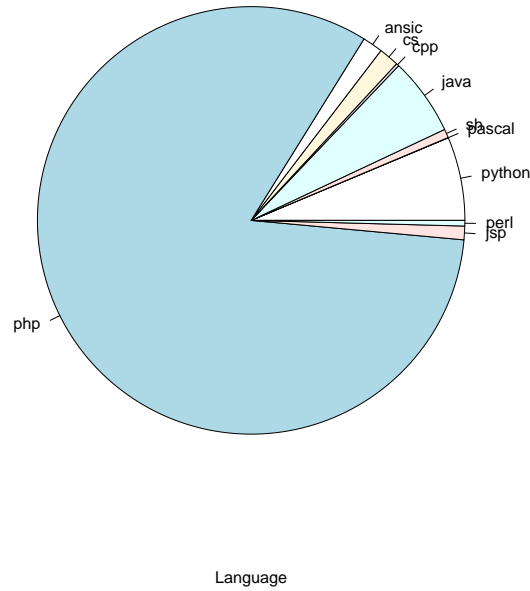


Figure 3.4: Programming languages used in Avoir forge

### 3.4.2 Knowledge Tree forge

Table 3.9 and figure 3.5 show results from applying the basic COCOMO model for cost estimation, to releases obtained from Knowledge Tree forge.

Project name	No. of SLOC	Effort	Schedule	No. developers	Estimated cost
ktdbipp	34,275	8.18 (98.16)	1.19 (14.28)	6.87	\$ 1,105,014
vao	3,702	0.79 (9.49)	0.49 (5.88)	1.61	\$ 106,782
ifilterindexer	328	0.06 (0.74)	0.19 (2.23)	0.33	\$ 8,381
securitydiagnos	316	0.06 (0.72)	0.18 (2.20)	0.33	\$ 8,060
ktwss	238	0.04 (0.53)	0.16 (1.97)	0.27	\$ 5,985

Table 3.9: Basic COCOMO model applied (using SLOCCount) to the largest Knowledge Tree forge software releases

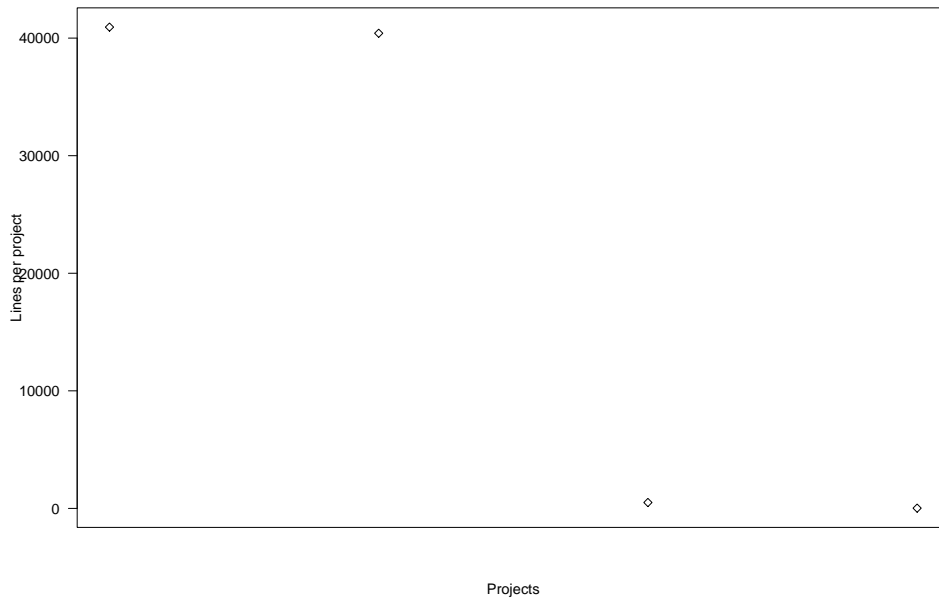


Figure 3.5: Lines of code in Knowledge Tree forge

Also, table 3.10 and figure 3.6 contain information regarding to the percentage of use of different programming languages in the Knowledge Tree forge.

<b>Programming language</b>	<b>No. of lines</b>	<b>Percentage</b>
PHP	40413	98.7223
C #	502	1.2263
SH	21	0.0512
Total	40936	100.0

Table 3.10: Programming languages used in Knowledge Tree forge

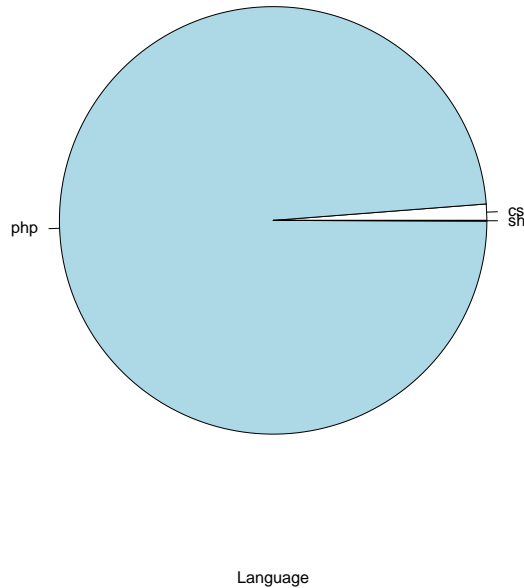


Figure 3.6: Programming languages used in Knowledge Tree forge

### 3.4.3 Aggregated data

In South African forges there is data from 54 projects obtained by SLOCCount. Table 3.11 shows a list with the projects with the highest developer/year effort relation. “nextgen”, “exe” and “zinf” projects are hosted in Avoir forge and “ktdbipp” and “vao” projects are hosted in Knowledge Tree forge.

Project name	No. of SLOC	Effort	Schedule	No. developers	Estimated cost
nextgen	164,634	42.50 (509.98)	2.23 (26.72)	19.09	\$ 5,740,994
ktdbipp	34,275	8.18 (98.16)	1.19 (14.28)	6.87	\$ 1,105,014
exe	15,304	3.51 (42.10)	0.86 (10.36)	4.07	\$ 473,900
zinf	13,873	3.16 (37.97)	0.83 (9.96)	3.81	\$ 427,485
vao	3,702	0.79 (9.49)	0.49 (5.88)	1.61	\$ 106,782

Table 3.11: Projects with higher effort

Also, table 3.12 contains information regarding the percentage of use of programming languages in the South African forges.

Programming language	No. of lines	Percentage
php	207925	85.1377
python	12837	5.2562
java	11964	4.8988
cs	3603	1.4752
ansic	3113	1.2746
jsp	2092	0.8565
sh	1358	0.5560
perl	862	0.3529
cpp	424	0.1736
pascal	44	0.0180
Total	244222	100.0

Table 3.12: Programming languages used in South African forges

### 3.5 Mailing Lists

There is a total of 14 mailing lists, and all of them have been analyzed and found in one South African forge: the Avoir forge (<http://gforge2.uwc.ac.za/index.php>).

Most of this mailing lists refer to empty mailing lists. Mailing lists which have not been created (but there is a link in the project's page) and their numerous commits mailing lists.

Local Forge	N. Mailing Lists	N. Posters	N. Messages
<a href="http://avoir.uwc.ac.za/">http://avoir.uwc.ac.za/</a>	14	0	0

Table 3.13: General mailing lists summary of results

In the Knowledge Tree forge, the Spider tool did not find any mailing lists. In order to validate this absence of mailing list, some random projects have been analyzed by hand. The results have shown that there are not mailing lists in this forge. For instance, some URL's have been visited such as WinSOO project<sup>5</sup> or VAO project<sup>6</sup> where mailing lists have not been activated yet<sup>7</sup>.

### 3.6 Authorship information

There are different type of authors in each project. Some of them, work for companies that are interested in participating in the project. The results in this section have been obtained from software releases (not from SCM repositories), by analysing copyright attributions in source files. Table 3.14 and figure 3.7 show these results.

<sup>5</sup>[http://forge.knowledgetree.com/mail/?group\\\_id=62](http://forge.knowledgetree.com/mail/?group\_id=62)

<sup>6</sup>[http://forge.knowledgetree.com/mail/?group\\\_id=58](http://forge.knowledgetree.com/mail/?group\_id=58)

<sup>7</sup>at 27th of March of 2007

Type of author	Detected number
Individual author	52
Teams and groups	2
Universities	4
Enterprises	9
Foundation or public entities	1
Anonymous/Unknown	3

Table 3.14: Authorship results

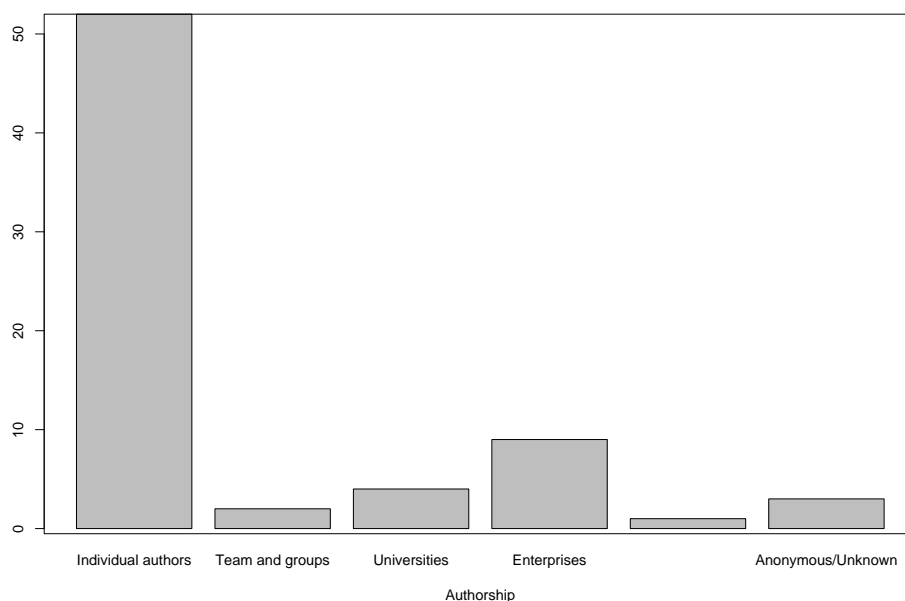


Figure 3.7: Authorship data (in software releases).

### 3.7 Sourceforge Analysis

Using the methodology outlined in the introduction of this report, several SourceForge developers have been identified as South African. This section includes data from the analysis of the activity of these developers in SourceForge.

Country	No. of committers	No. of commits
South Africa	102	53142

Table 3.15: General results from Sourceforge

Table 3.16 shows committers with the highest number of commits:

Project name	Committer	No. of commits
turbocash	delphidreamer	18562
turbocash	sylvainwn	9555
underpressure	wolfman_za	2737
coefficient	detkin	1919
cilib	espeer	1752

Table 3.16: The most active South African committers detected in Sourceforge

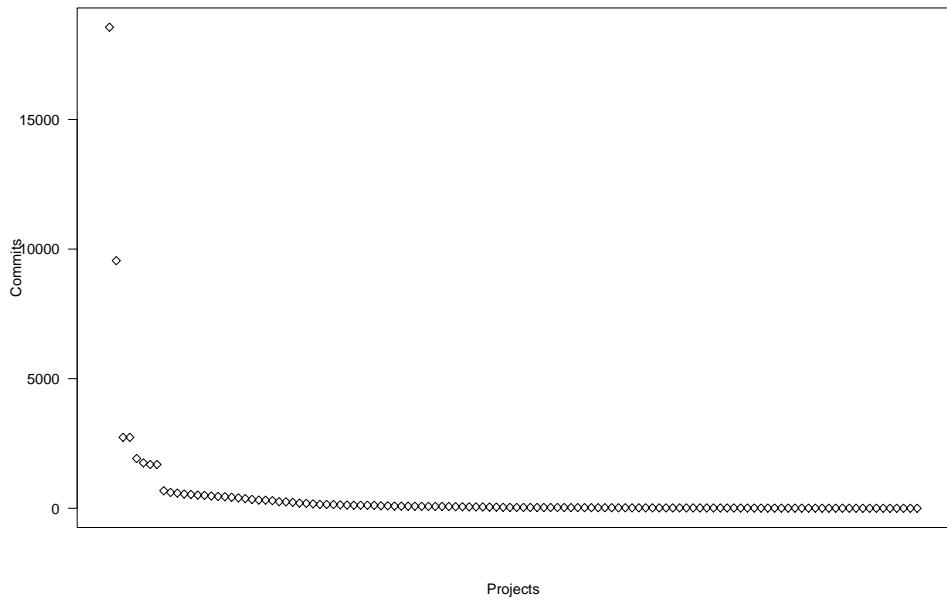


Figure 3.8: Commits in SourceForge

## Chapter 4

# Appendixes

### 4.1 General information

The following is a list of tables showing a relation of information provided by our partners.

#### Communities

Table 4.1: Communities detected by partners

Name	URL
FSIU	<a href="http://fsiu.uwc.ac.za/">http://fsiu.uwc.ac.za/</a>
Geek Freedom League	<a href="http://www.freedomleague.org.za/">http://www.freedomleague.org.za/</a>
IPA	<a href="http://www.lpa.org.za/">http://www.lpa.org.za/</a>

#### LUGs

Table 4.2: South African LUGs

LUG Name	URL
Linus groups	<a href="http://www.linux.org/groups/southafrica.html">http://www.linux.org/groups/southafrica.html</a>
Several Linux user groups in Africa	<a href="http://www.tectonic.co.za/">http://www.tectonic.co.za/</a>
News portal	<a href="http://www.fossfa.net/">http://www.fossfa.net/</a>

#### Projects

Table 4.3: Projects collected by partners

<b>Project name</b>	<b>URL</b>
AssetScout	<a href="http://www.arkateq.com">http://www.arkateq.com</a>
Bika lab systems	<a href="http://www.bika.co.za">http://www.bika.co.za</a>
Central Case Management System	<a href="http://www.ccms.org.za">http://www.ccms.org.za</a>
eLearn	<a href="http://www.e81.co.za">http://www.e81.co.za</a>
Employee Productivity Suite	<a href="http://www.softwarefutures.com">http://www.softwarefutures.com</a>
Enterprise Resource, Customer Management System (ERCUMS)	<a href="http://www.arkateq.com">http://www.arkateq.com</a>
Firestorm	<a href="http://www.ku-shan.co.za/firestorm">http://www.ku-shan.co.za/firestorm</a>
Impi Linux	<a href="http://www.impi.org.za">http://www.impi.org.za</a>
MobileED	<a href="http://mobiled.uiah.fi/">http://mobiled.uiah.fi/</a>
NetServant	<a href="http://www.arkateq.com/index.php?ID=13">http://www.arkateq.com/index.php?ID=13</a>
OpenOffice Pro	<a href="http://www.openofficepro.org">http://www.openofficepro.org</a>
OpenOffice.org	<a href="http://www.openoffice.org.za/">http://www.openoffice.org.za/</a>
OpenOffice.org Association of South Africa	<a href="http://www.openoffice.org.za/oasa">http://www.openoffice.org.za/oasa</a>
Opensource Network Management Tools	<a href="http://www.excubedconsulting.com">http://www.excubedconsulting.com</a>
OSS Excellency Center	<a href="http://www.novell.co.za">http://www.novell.co.za</a>
Shuttleworth Foundation Website CMS	<a href="http://www.liquidthought.co.za">http://www.liquidthought.co.za</a>
SL4 - Lotto database application	<a href="http://www.slx.za.net">http://www.slx.za.net</a>
Translate.org.za	<a href="http://www.translate.org.za">http://www.translate.org.za</a>
TTS-Cubed	<a href="http://sourceforge.net/projects/tts-cubed/">http://sourceforge.net/projects/tts-cubed/</a>
Vector Data Language	<a href="http://www.e81.co.za">http://www.e81.co.za</a>
www.viva2010.co.za	<a href="http://www.viva2010.co.za">http://www.viva2010.co.za</a>

### Platforms/Forges

Table 4.4: Platforms collected by partners

<b>Description</b>	<b>URL</b>
Government	<a href="http://www.oss.gov.za/">http://www.oss.gov.za/</a>
National Advirosy Council on Innovation	<a href="http://www.naci.org.za/floss/">http://www.naci.org.za/floss/</a>
Global Best Practice in Free Software Adoption	<a href="http://wiki.go-opensource.org/taskforce/">http://wiki.go-opensource.org/taskforce/</a>
Technological issues	<a href="http://www.tectonic.co.za/searchresults.php">http://www.tectonic.co.za/searchresults.php</a>
FLOSS projects	<a href="http://www.go-opensource.org/mosaddphp/gossip/project_all.php">http://www.go-opensource.org/mosaddphp/gossip/project_all.php</a>
Organization	<a href="http://www.linux.co.za/">http://www.linux.co.za/</a>
Organization	<a href="http://www.raspberry.co.za/raspberry/">http://www.raspberry.co.za/raspberry/</a>

Continued on next page

Table 4.4: Platforms collected by partners

Description	URL
South African founder	<a href="http://www.ubuntu.com/">http://www.ubuntu.com/</a>
South African Ubuntu localization	<a href="http://www.impi.org.za/">http://www.impi.org.za/</a>
SA distro, originally educational	<a href="http://www.getopenlab.com">http://www.getopenlab.com</a>
Training and LPI certification	<a href="http://www.ledge.co.za/lpi/index.php">http://www.ledge.co.za/lpi/index.php</a>
FLOSS in schools	<a href="http://www.tuxlabs.org.za/">http://www.tuxlabs.org.za/</a>
FLOSS in schools	<a href="http://www.inkululeku.co.za/">http://www.inkululeku.co.za/</a>
FLOSS in schools	<a href="http://www.netday.org.za/">http://www.netday.org.za/</a>

## 4.2 Data from repositories

The followings are tables showing data from the repository mining with CVSanaly in Avoir and SourceForge forges.

Table 4.5: SCM results for Avoir Forge

Project name	Committer	Commits
nextgen	tohir	9679
nextgen	dkeats	7186
odimex	dkeats	6712
nextgen	joconnor	5060
nextgen	megan	4333
nextgen	cadrian	4085
odimex	tohir	3562
nextgen	wnitsckie	2770
nextgen	kevinc	2501
odimex	wnitsckie	2458
odimex	joconnor	2415
nextgen	jsc	2177
sems	hamza	1856
nextgen	jonathan	1610
nextgen	paulscott	1601
chisimba_skins	adrian	1584
odimex	megan	1532
zinf	justin	1480
odimex	jsc	1413
nextgen	slegassick	1293
odimex	slegassick	1283
sems	sergem	1281
odimex	jonathan	1231
nextgen	morpheus	1053

Continued on next page

Table 4.5: SCM results for Avoir Forge

<b>Project name</b>	<b>Committer</b>	<b>Commits</b>
chisimba_skins	tohir	1036
chisimba_skins	wnitsckie	990
sems	ghinde	925
nextgen	wwindvogel	887
nextgen	ghinde	866
odimex	paulscott	859
sems	colleen	820
nextgen	prince	805
sems	jadam	745
sems	ameyer	676
nextgen	nic	618
chisimba_skins	nic	610
voice_jain_sip	jbangoo	601
chisimba_skins	cadrian	569
sems	joconnor	556
softbridge	jbangoo	554
realtime	adrian	552
nextgen	jadam	536
nextgen	sola	534
nextgen	karitz	509
xultools	pmbekwa	448
nextgen	cswanson	447
chisimba_skins	paulscott	445
freemind_src_0.7.1	wnitsckie	432
cshe	megan	396
nextgen	bishar	356
odimex	prince	343
zinf	lretief	343
nextgen	mapijola	300
sems	wwindvogel	297
nextgen	osamuel	286
nextgen	sergem	243
odimex	smtegha	233
nextgen	smtegha	233
odimex	wangzy	209
nextgen	wangzy	209
chisimba_cms	pmbekwa	173
odimex	karitz	170
nextgen	ftilu	159
nextgen	abakpaj	150
tabeisa	tohir	137
nextgen	liesel	134
nextgen	chellmuth	121
chisimba_cms	megan	112

Continued on next page

Table 4.5: SCM results for Avoir Forge

Project name	Committer	Commits
xultools	paulscott	112
realtime	davidwaf	105
nextgen	ameyer	101
nextgen	deanvn	100
nextgen	swentzel	99
nextgen	jbeneke	90
nextgen	abdurahim	88
nextgen	hamza	83
smarty_module	dkeats	77
nextgen	davidwaf	65
chisimba_skins	dkeats	56
odimex	osamuel	56
nextgen	james	50
odimex	james	50
postgis	istamp	48
documentation	dkeats	47
pbl	fernando	46
odimex	nina	45
nextgen	nina	45
realtime	mohamed	43
nextgen	nsabagwa	42
odimex	ftilu	40
postgis	paulscott	40
nextgen	irshaad	37
chisimba_skins	megan	33
chisimba_skins	kevinc	33
chisimba_cms	paulscott	33
nextgen	kudakwashe	29
nextgen	ldomingos	28
odimex	abakpaj	26
nextgen	mfall	20
documentation	tohir	19
nextgen	ikaddu	18
documentation	paulscott	16
nextgen	pkuti	15
odimex	ldomingos	13
tabeisa	kevinc	13
sems	askippers	13
nextgen	colleen	11
sems	jsc	11
tabeisa	jsc	10
simba	jsc	10
phpcart	jsc	10
chisimba_skins	philipp	10

Continued on next page

Table 4.5: SCM results for Avoir Forge

<b>Project name</b>	<b>Committer</b>	<b>Commits</b>
softbridge	jsc	10
linuxradio	jsc	10
nextgen	mpscheidt	9
odimex	davidwaf	9
odimex	joselee	9
nextgen	acej	9
zinf	caron	9
nextgen	joselee	9
documentation	jmulindwa	8
documentation	rwhitney	7
chisimba_skins	jsc	6
nextgen	snyakaisiki	6
documentation	deanvn	6
chisimba_skins	hamza	5
chisimba_skins	jadam	4
odimex	kesun	4
nextgen	kesun	4
chisimba_skins	rwhitney	4
chisimba_skins	joconnor	3
cshe	nic	3
nextgen	fernando	3
postgis	martin_weis	3
chisimba_skins	prince	3
odimex	fernando	3
documentation	nic	3
chisimba_skins	ghinde	3
nextgen	askippers	2
realtime	jsc	2
documentation	nsabagwa	2
chisimba_skins	colleen	2
odimex	mfall	2
documentation	adrian	2
chisimba_skins	ameyer	2
nextgen	fakorli	1
documentation	lilji25	1
freecourseware	philipp	1
chisimba_skins	sergem	1
odimex	pentech	1
realtime	jbangoo	1
nextgen	pentech	1

Table 4.6: SCM results in SourceForge forge

<b>Project name</b>	<b>Committer</b>	<b>Commits</b>
turbocash	delphidreamer	18562
turbocash	sylvainwn	9555
underpressure	wolfman.za	2737
underpressure	wolfman.za	2737
coefficient	detkin	1919
cilib	espeer	1752
sdlpure	entro-p	1689
sdlpure	entro-p	1689
cricketsb	adriaan	677
ieee1394diag	mztse	612
afrikaans	ysbeer	587
dump3	pyropunk	551
cricketsb	chappies	533
be4gle	dataphile	511
knab	vhata	498
iscan	ripienaar	475
garchiver	life	456
frags	i.am.the.user	446
xode	everth	427
impi	gvandyk	401
qmc	mikew	374
basn	colinwebber	341
coefficient	lbutgereit	319
rot13	frith88	314
angua	bshuttleworth	295
coefficient	pvzyl	255
dgp	alanpowell	250
cilib	gpampara	231
glasscheckers	wduminy	202
abigail	adejongh	190
oschooladmin	avrin17	170
garchiver	droux	149
binarypatch	wolfman.za	147
qmc	michael	145
masqserver	traindrv_	134
collect	ballestr	126
logfiles	andret	119
s3fc	alberts3	117
afid	nicross	117
fpkg	jlea	114
accessserver	thaig	97
mozarella	perspex	96
gdcsim	berndfoobar	85
turbocash	ettiend	83

Continued on next page

Table 4.6: SCM results in SourceForge forge

Project name	Committer	Commits
tnghttpd	encrypted_	81
garchiver	znh	79
coefficient	richardwatson	76
coefficient	tfogwill	74
carmanager	kayonlord	73
afid	jodilewis	72
gloworm	davidturner	69
pabxmon	pranesh	63
multidb	deneb	60
openrecbook	terencevs	57
idms	nkukard	55
slfm	kodgehopper	54
ftdifullspddrv	sa_lunatic	48
collect	sash	47
makeme	jewel	38
redtea	jewel	35
pbof	gvandyk	34
telpos	cyan	34
zmessagequeue	elabuschagne	32
zmessagequeue	elabuschagne	32
abigail	rviljoen	31
pyado	davidfraser	31
eatrss	thinguy	31
cricketsb	anirban_c8	29
pypaint	igalboc	28
afid	allandewar	27
feclib	nroets	26
agere	grobh	25
cilib	clive9900855	25
cilib	jkroon	25
collect	uid53429	23
garchiver	jaco	22
osfac	osfac	22
mydbphp	keywest	20
abigail	tboonzaaier	20
eduslug	philipajane	19
kolabwebclient	chappies	19
glpoisson	nroets	18
slaughter	entro-p	18
cilib	apengelbrecht	18
slaughter	entro-p	18
garchiver	raven	17
s3fc	mcurf	17
collect	sb	16

Continued on next page

Table 4.6: SCM results in SourceForge forge

Project name	Committer	Commits
nhtexttile	entro-p	16
pycontrol	andrewfraser	15
money	mwdiablo	14
cilib	fvnkerk	14
garchiver	uid43216	14
meetmin	xwriter	11
iscan	uid38242	11
cricketsb	uid501	10
eduslug	soloffice	8
agere	stewartallen	8
modjldap	joevdberg	8
multidb	bain	7
afid	lucask	7
freeview	nobbyknox	7
puxedo	ikklit	5
puxedo	ikklit	5
pypaint	davidfraser	4
garchiver	phoenix	4
phpfileutils	moonbag	4
gluetils	bluecypher	4
chronologer	nroets	4
gluetils	bluecypher	4
newsclass	hbro	3
openufs	dewitpj	2
cilib	hmkunzmann	2
impi	andrelouw	2
afrikaans	neaj	2
openrecbook	rygorka	1
kolabwebclient	wolfman.za	1
afrikaans	droux	1
precis	colinwebber	1
bindump	mzito	1

### 4.3 COCOMO model applied to software releases

The following is a list of tables showing data obtained after applying basic COCOMO model to projects found in the forges.

#### Avoir forge

Table 4.7: Basic COCOMO model applied to Avoir software distributions

Project name	SLOC	EPM(EPY)	Est. Years	Est Dev.	Est. Cost
nextgen	164,634	42.50 (509.98)	2.23 (26.72)	19.09	\$ 5,740,994
exe	15,304	3.51 (42.10)	0.86 (10.36)	4.07	\$ 473,900
zinf	13,873	3.16 (37.97)	0.83 (9.96)	3.81	\$ 427,485
c2go	3,243	0.69 (8.25)	0.46 (5.58)	1.48	\$ 92,926
ziggis	3,101	0.66 (7.88)	0.46 (5.48)	1.44	\$ 88,658
mailer	1,704	0.35 (4.20)	0.36 (4.31)	0.97	\$ 47,281
postgis	1,233	0.25 (2.99)	0.32 (3.79)	0.79	\$ 33,663
phpeditini	194	0.04 (0.43)	0.15 (1.81)	0.24	\$ 4,829

### Knowledge tree forge

Table 4.8: Basic COCOMO model applied to Knowledge Tree forge software distributions

Project name	SLOC	EPM(EPY)	Est. Years	Est Dev.	Est. Cost
ktdbipp	34,275	8.18 (98.16)	1.19 (14.28)	6.87	\$ 1,105,014
vao	3,702	0.79 (9.49)	0.49 (5.88)	1.61	\$ 106,782
cmdbulkimp	520	0.10 (1.21)	0.22 (2.69)	0.45	\$ 13,597
quicklinks	349	0.07 (0.79)	0.19 (2.29)	0.35	\$ 8,946
ifilterindexer	328	0.06 (0.74)	0.19 (2.23)	0.33	\$ 8,381
securitydiagnos	316	0.06 (0.72)	0.18 (2.20)	0.33	\$ 8,060
ktwss	238	0.04 (0.53)	0.16 (1.97)	0.27	\$ 5,985
newsdashlet	212	0.04 (0.47)	0.16 (1.88)	0.25	\$ 5,300
documentfolderl	201	0.04 (0.45)	0.15 (1.84)	0.24	\$ 5,012
inddiagnostic	162	0.03 (0.35)	0.14 (1.69)	0.21	\$ 3,996
repositorystats	140	0.03 (0.30)	0.13 (1.59)	0.19	\$ 3,428
foldertree	54	0.01 (0.11)	0.09 (1.09)	0.10	\$ 1,261
imageindexer	50	0.01 (0.10)	0.09 (1.06)	0.10	\$ 1,163
ktclamav	47	0.01 (0.10)	0.09 (1.03)	0.09	\$ 1,090
moreviewcolumns	45	0.01 (0.09)	0.08 (1.01)	0.09	\$ 1,041
docdetaillinks	32	0.01 (0.06)	0.07 (0.88)	0.07	\$ 728
xmlhtml	29	0.00 (0.06)	0.07 (0.85)	0.07	\$ 656
kt-slovak	21	0.00 (0.04)	0.06 (0.75)	0.06	\$ 468
compactmetadata	19	0.00 (0.04)	0.06 (0.72)	0.05	\$ 421
czechplugin	17	0.00 (0.03)	0.06 (0.69)	0.05	\$ 375
ktgerman	15	0.00 (0.03)	0.05 (0.65)	0.04	\$ 329
condmetadata	15	0.00 (0.03)	0.05 (0.65)	0.04	\$ 329
hunlangpack	15	0.00 (0.03)	0.05 (0.65)	0.04	\$ 329
kt3italian	11	0.00 (0.02)	0.05 (0.58)	0.04	\$ 237
slo-pack	11	0.00 (0.02)	0.05 (0.58)	0.04	\$ 237

Continued on next page

Table 4.8: Basic COCOMO model applied to Knowledge Tree forge software distributions

<b>Project name</b>	<b>SLOC</b>	<b>EPM(EPY)</b>	<b>Est. Years</b>	<b>Est Dev.</b>	<b>Est. Cost</b>
scmsvn	11	0.00 (0.02)	0.05 (0.58)	0.04	\$ 237
ktpolish	11	0.00 (0.02)	0.05 (0.58)	0.04	\$ 237
ktitalian	11	0.00 (0.02)	0.05 (0.58)	0.04	\$ 237
smpl-chn	10	0.00 (0.02)	0.05 (0.56)	0.03	\$ 215
ktrus	10	0.00 (0.02)	0.05 (0.56)	0.03	\$ 215
catalanplugin	10	0.00 (0.02)	0.05 (0.56)	0.03	\$ 215
spanish-co-302	10	0.00 (0.02)	0.05 (0.56)	0.03	\$ 215
frenchplugin	10	0.00 (0.02)	0.05 (0.56)	0.03	\$ 215
kt-swedish	10	0.00 (0.02)	0.05 (0.56)	0.03	\$ 215
japanese	10	0.00 (0.02)	0.05 (0.56)	0.03	\$ 215
dutch	9	0.00 (0.02)	0.04 (0.53)	0.03	\$ 192
pt-br					
spainplugin					