The Business of Free Software: Enterprise Incentives, Investment, and Motivation in the Open Source Community

Dr. Marco Iansiti, Ph.D.
Gregory L. Richards

Marco Iansiti

David Sarnoff Professor of Business Administration

Harvard Business School

Gregory L. Richards

Managing Director

Keystone Strategy, Inc.

_____

The Business of Free Software:

Enterprise Incentives, Investment, and Motivation in the Open Source Community

_____

Marco Iansiti, David Sarnoff Professor of Business Administration, Harvard Business School
Gregory L. Richards, Managing Director, Keystone Strategy, Inc.

Boston, MA 02163, USA
October 9, 2006

PRELIMINARY DRAFT – COMMENTS WELCOME

ABSTRACT

In this paper, we examine the motivations of large information technology ("IT") vendors, to invest in open source software ("OSS"). What drives companies with large, proprietary software portfolios to invest hundreds of millions of dollars in OSS? We approach this question by grouping a sample of OSS projects into clusters and examining vendors' motivations for each cluster. We find one cluster has received almost no investment. Contributions to projects in this cluster are confined to the voluntary effort of the vendors' employees, and vendors are likely altruistically motivated. By contrast, the other cluster has received over 99% of vendor investments. Here, vendors are more likely economically motivated to invest in OSS projects that can serve as a complementary asset to vendors' core, proprietary businesses.

Keywords: OSS; Open-Source; IT Vendor Strategies

INTRODUCTION

The increasing profile and influence of open source software (OSS) has been matched by a growing body of literature seeking to explain the motivations of contributors[2,3,4,5] Most research has focused on individual contributors[6,7,8,9,10], even though IT vendors have made significant contributions. IBM has contributed more than $1 billion to the development and promotion of the Linux operating system[11], and other vendors such as Sun are ramping up OSS efforts and investment.

In this paper, we examine the motivations of large companies, information technology ("IT") vendors, to invest in OSS. Why do IT vendors that have historically sold proprietary, "closed source" software invest millions of dollars in OSS? Where have they chosen to invest, and what are the characteristics of the OSS projects they contribute to?

Previous work on motivations of contributors has examined either a small number of differentiated OSS projects[12] or a multitude of OSS projects as a single homogenous collection[13]. Our approach falls between these two extremes: we group OSS projects into clusters and analyze IT vendors' motivations by cluster. The clusters are delineated by the level of IT vendor investment, and by the impact of the application, indicated by the application installation rate (defined as the number of weekly distributions of that software times the size of the installation file). Clustering by these two variables is an objective way to identify the group of OSS projects which have received investments and made an impact. Examining the nature of these projects sheds light on vendors' motivations for investing in this group.

We find that Cluster 1 consists of projects that have received almost all of vendor investments – we refer to these as the "money-driven cluster". We find that Cluster 2 consists instead of a large number of projects that have received almost no vendor investment – we refer to these as the "community-driven cluster". We estimate that the eighteen "money-driven cluster" projects have received over $2 Billion dollars in investment. On the other hand, IT vendors have generally ignored projects in the "community-driven cluster" and appear to lack a coordinated strategy in dealing with them. The only support appears to be the strictly voluntary effort of the vendors' employees. Finally, examining the impact of projects in both clusters showed that vendors have not invested uniformly in high impact projects. Rather, it appears that vendors are investing in high impact OSS projects that can to serve in a complementary fashion to draw revenues to their own (largely proprietary) core businesses.

This paper is organized as follows. First, we briefly review relevant literature on individual motivations, corporate motivations, and classifications of OSS projects. We then describe the clustering approach and OSS project sample for the paper. Next, we report findings on OSS project clustering. We follow with a discussion of vendor motivations in each cluster, the benefits accrued to vendors supporting OSS project, and the strategies vendors employ. We conclude by outlining directions for future research.

BACKGROUND

In this section, we review previous work on motivations for individuals and companies to contribute to or support OSS projects. We also discuss previous methods and metrics used to classify OSS projects

Past research characterized the motivations to contribute to OSS software as either intrinsic (i.e., the contribution is valued for its own sake)[14,15] or extrinsic (i.e., the contribution provides indirect benefits)[16,17]. Two examples of intrinsic motivations are intellectual stimulation and philosophical beliefs in software being open.[18] Two examples of extrinsic motivations are the "signaling incentives"[19] of career concerns[20] and reputation benefits. Research has found individuals are not driven exclusively by intrinsic or extrinsic motivations but by a combination of the two types of motivations. Ghosh et al. surveyed over 2,700 developers[21] about the principal reasons they joined the OSS community. Almost 80% of respondents said it was to learn new knowledge and skills, and 50% mentioned sharing knowledge and skills – two intrinsic motivations. At the same time, over 50% mentioned directly (i.e., getting paid for working on OSS) or indirectly (e.g., improving job opportunities) earning money, and getting help developing a new software product – two extrinsic motivations.

Examining motivations for IT vendors to contribute to OSS projects adds insight to an area described as unaddressed[22] and not well understood[23]. Rossi and Bonaccorsi adapted the open source motivational frameworks for individuals to analyze vendors supplying open source-based products and services. The intrinsic individual motivations of "sense of belonging to a community" and "fighting against proprietary software" became altruistic vendor motivations of "not betraying the developer community's trust"

6

and "reducing the market power of large software companies". The extrinsic individual

motivations of "monetary rewards" or "filling an unfilled market" became economic

motivation of "the ability to supply [open source] software-related services". Rossi and

Bonaccorsi's survey of 146 Italian vendors producing open source solutions found 19%

were altruistic in their motivations and practices, 34% were economic in both motivations

and practices, and the rest exhibited some mixture of altruistic and economic motivations

and practices.[24] They found vendors' motivations were a mix of intrinsic and extrinsic,

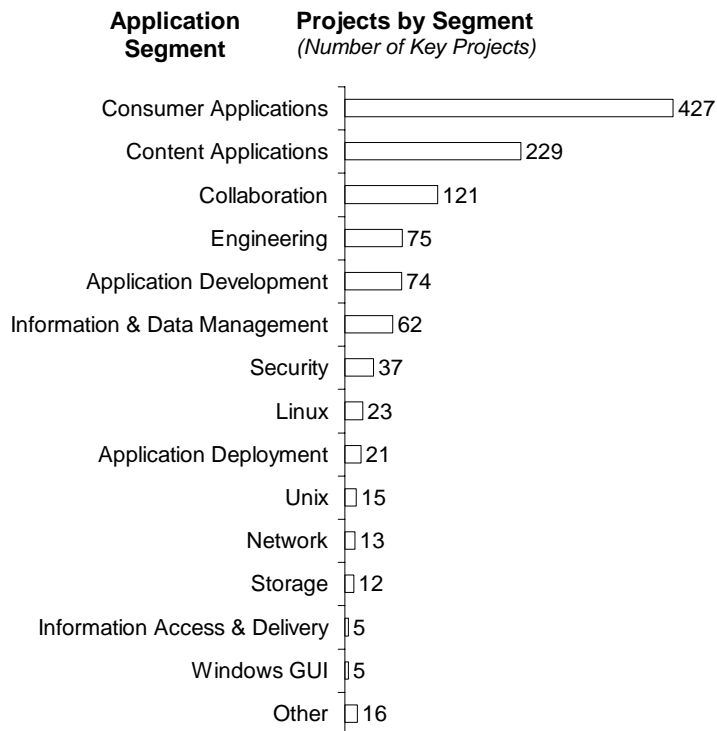similar to individuals' motivations.

Moving from motivations to classifications, we review Nakakoji et al.'s

classification of OSS projects. They examined the different collaboration models used in

various projects and classified OSS projects in three ways: exploration-oriented, utility-

oriented, and service-oriented.[25] Exploration-oriented projects focus on pushing the front

line of software development, have high code quality, and are driven by a lead

programmer. Utility-oriented projects focus on filling voids in functionality and have a

modular and independent sub-project structure with little centralized control. Service-

oriented projects focus on providing stable and robust services to all OSS stakeholders

through conservative incremental improvement and are led by a council. For each type

of OSS project, this approach provided a valuable means of identifying organizational

styles and management techniques best-suited to a improving the probability of

successfully developing the desired piece of OSS software. However, as the authors

noted, this categorization can not be applied to all OSS projects and does not evaluate a

project's impact.

A framework that can be applied to all OSS projects requires common metrics. Previously developed metrics have measured the success of OSS projects. Crowston et al. adapted standard measurements of success used in information systems research to OSS efforts. The result was several indicators of system and information quality (e.g., user satisfaction, use, and individual and organizational impact).[26] Their framework was useful in helping the academic and commercial worlds grasp the evolution of some of the more high-profile and well-publicized OSS projects. However, their metrics depend on personal opinion and subjective ratings. The measure of OSS impact used in this paper aims to address problems of subjectivity by employing measurable objective metrics which can be universally applied to all OSS projects.

METHODOLOGY

To perform our analysis, we first assemble a sample of OSS projects. The OSS project sample consists of projects with community interest and significant activity. We use the list of OSS projects categorized and profiled on www.wikipedia.org as an initial sample source because it is developed and maintained by the OSS community. As a result, it is a reasonable representation of what the community collectively considers to be the most important projects. At the start of the study period (June 2006), www.wikipedia.org listed 1,135 active OSS projects. The projects are categorized by software segment, and Figure 1 presents the number of OSS projects in each major segment.

Figure 1: OSS Projects by Segment[27]

Application Segment	Projects by Segment (Number of Key Projects)
Consumer Applications	427
Content Applications	229
Collaboration	121
Engineering	75
Application Development	74
Information & Data Management	62
Security	37
Linux	23
Application Deployment	21
Unix	15
Network	13
Storage	12
Information Access & Delivery	5
Windows GUI	5
Other	16

Next, we narrowed the list of projects to those with significant activity, with the cut-off defined (somewhat arbitrarily) as projects with community involvement equivalent to a full-time, seven person development team. The average time per participant was calculated by using the distribution of time spent found in the European FLOSS survey, which is reproduced in Table 1.

Table 1: Distribution of Time Spent on OSS Development[28]

Amount of time per week	Percent of Survey Respondents
< 2 hours	22.5%
2 - 5 hours	26.1%
6 - 10 hours	20.9%
11 - 20 hours	14.3%
21 - 40 hours	9.1%
> 40 hours	7.1%

To calculate a weighted average, we use the midpoint of each range as the average value, and assume that no one works more than 60 hours per week. I.e., we assume the 22.5% of participants who spend less than 2 hours per week on OSS development spend an average of 1 hour, and the 7.1% of participants who spend more than 40 hours per week on OSS development spend an average of 50 hours. The resulting weighted average is 11.4 hours per week. Consequently, 26 OSS participants correspond to the equivalent of a 7 FTE team. Accordingly, the Wikipedia projects were filtered for being active with troves on SourceForge, having 26 or more developers involved, and having downloads in 2005.

Since some well-known projects (e.g., Linux) are not downloaded from SourceForge, 9 obviously important projects were added to the sample set: Linux, Firefox, OpenOffice, MySQL, SugarCRM, JBoss, PHP, Perl, and Python. The final sample set consisted of 50 OSS projects.

Previous OSS metrics have been based on classic information system success measures of project quality, user satisfaction, use, individual impact (e.g., productivity, rate of learning), and organizational impact (e.g., responsiveness, coordination, quality of decision-making).[29] Others have extended this set of metrics to include project output, activity and outcomes for project members.[30] These kinds of metrics are well suited for looking at individual projects, but the subjective nature of some metrics (e.g., user satisfaction) makes it difficult to apply them consistently across projects. A more appropriate metric for looking across projects is a quantitative, more objective one, such as the number of downloads, which has been often used.[31] However, looking at downloads does not account for the complexity of the problem solved by the OSS project. Based on these considerations, we used the metric of installation rate, which is defined as weekly distributions of the software times the size of the installation file. The number of distributions is a measure of demand or reach, and installation file size is a proxy for problem complexity and programming effort required. Multiplying weekly distributions and file sizes results in a quantitative indicator of the impact of an OSS project. For some segments (e.g., file-sharing, security), we recognize reach or demand may be overstated.

All data was collected through publicly available information sources and third-party analyst reports[32]. Large IT Vendors' public contributions to OSS Projects in each

category were identified through extensive press searches of company press releases about specific OSS projects, OSS, general IT & business publications, and announcements from the individual project web sites. IT vendors' annual reports provided the breakdown of their 2005 revenues by hardware, software, and services. The individual sources for this information are cited below in the references for each Figure and Table.

Individual OSS projects' download volume, image size, and operating system data were collected from www.sourceforge.net and individual project Web sites. As Linux is distributed from several Web sites that do not report distributions, IDC's 2005 Linux installation data was used as a proxy[33].

Clusters of OSS projects were identified using both the Pearson Hierarchical Clustering and Spearman Rank Correlation techniques. The Pearson technique minimizes within-segment variation and maximizes between-segment variation. The Pearson process began by assigning each OSS project to a separate cluster. Pair-wise distances between clusters were then calculated using the Pearson similarity distance measure with Installation Rate and Vendor Investment as the variables. A between cluster distance matrix was then constructed using the distance values. Clusters were then sequentially paired and analyzed using the average between-group linkage method.[34] The Spearman process begins by ranking each project by variable (e.g., the project with the highest installation rate is given an installation rate ranking of 1). If more than one project has the same value for a variable, they are all given the average of the ranking places (e.g., if four projects shared ranks 5, 6, 7, and 8, they would be each be ranked

6.5). Pair-wise distances between clusters were then calculated using the Spearman

Correlation Coefficient as the distance measure.

ANALYSIS

Figure 2 is a plot of OSS projects' installation rates on a log scale. It shows installation rates were heavily skewed to the top four projects. The rates then fell off quickly, stabilized, and then fell again into a long tail. In Figure 2 only the top 20 OSS projects by installation rate are shown for clarity. The installation rates for Apache, MySQL, Perl, Python, and PHP may be undercounted because they were often shipped with Linux as part of the LAMP[35] stack web and application servers. Apache may be further undercounted as our data is based on a survey of publicly accessible web sites, and this methodology does not count private web servers like corporate intranets.

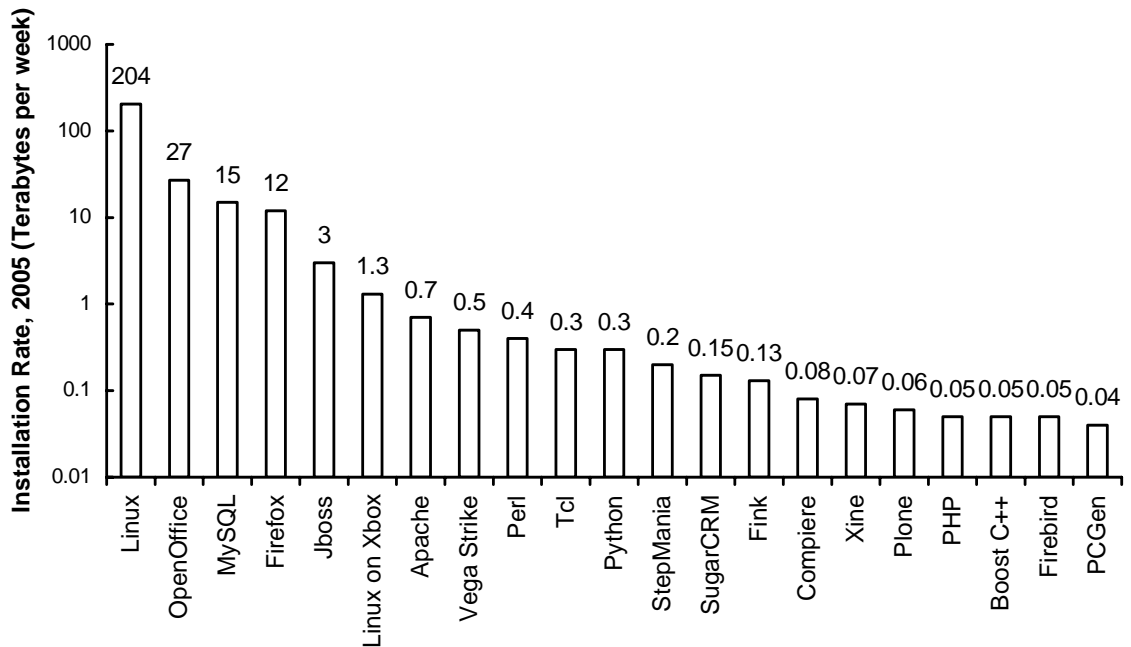**Figure 2: OSS Installation Rate Distribution (Log Plot)[36]**

Figure 3 is a logarithmic plot of the investments from 1995-2005 in the OSS projects. The investment is much more concentrated and falls off more sharply than the installation rate. Similar to Figure 2, only the projects with the top 20 investment rates are shown for clarity. Non-cash contributions are valued at the equivalent costs. E.g., if

developer manpower is dedicated to a project, then the fully loaded cost of the developers[37] is counted in the investment total. If previously developed code is open-sourced, and incorporated into an OSS project, then the cost of developing that code is counted towards the total. Venture capital investments in start-up companies are included because the venture capital firms are investing in companies that will become IT vendors. Venture capital investments are treated like cash contributions.

OSS projects without are considered to have received no investment if a literature search turns up no mention of an IT vendor or venture capital firm investing in the project, and the official project web sites listed on www.sourceforge.net or www.wikipedia.com do not mention a non-profit foundation associated with the project, or if communication with the project leader indicates no donations were received. This methodology understates the total investment in OSS by not including customer contributions (e.g., cash donations or resources spent creating additional code to deploy or customize an OSS project for their own use), volunteer time, or donations from individuals or other non-profit foundations. These sources are excluded since the focus of this paper is vendors' motivations.
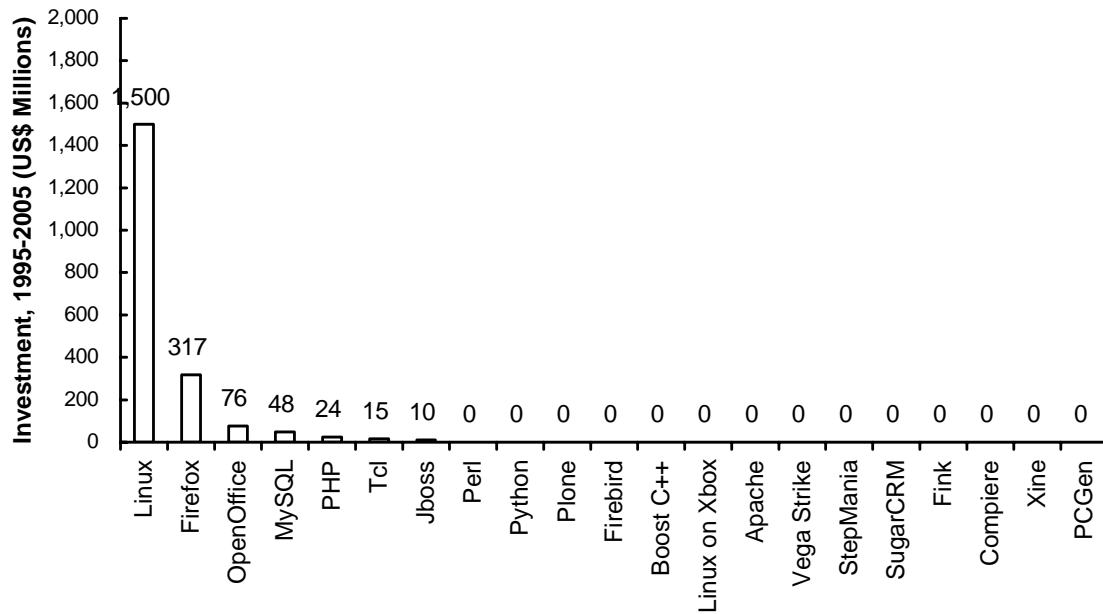
Figure 3: OSS Investment Distribution[38]

Investment, 1995-2005 (US$ Millions)

Project	Investment
Linux	1,500
Firefox	317
OpenOffice	76
MySQL	48
PHP	24
Tcl	15
Jboss	10
Perl	0
Python	0
Plone	0
Firebird	0
Boost C++	0
Linux on Xbox	0
Apache	0
Vega Strike	0
StepMania	0
SugarCRM	0
Fink	0
Compiere	0
Xine	0
PCGen	0

Table 2: Primary OSS Project Cluster by Clustering Method[39]

Pearson Correlation	Absolute Pearson Correlation	Spearman Rank Correlation
Linux	Linux	Linux
Firefox	Firefox	Firefox
Open Office	Open Office	Open Office
JBoss	JBoss	JBoss
Open Web Application Security Project (OWASP)	Open Web Application Security Project (OWASP)	Open Web Application Security Project (OWASP)
Tcl	Tcl	Tcl
Xemacs	Xemacs	Xemacs
PHP	PHP	PHP
SugarCRM	SugarCRM	SugarCRM
OpenBeos	OpenBeos	OpenBeos
Advanced Authoring Format (AAF)	Advanced Authoring Format (AAF)	Advanced Authoring Format (AAF)
MySQL	MySQL	MySQL
Firebird	Firebird	Firebird
Plone	Plone	Plone
Freenet	Freenet	Freenet
NSLU	NSLU	
Python	Python	
Perl	Perl	

Table 2 shows the primary OSS project cluster identified from the analysis. There are three projects that fall into the primary cluster under the Pearson method, but not the Spearman Rank method. To understand which classification is more appropriate, the projects were plotted by installation rate and investment, as shown in Figure 4, with guide lines drawn using the more inclusive clustering from the Pearson method. For clarity, not all of the projects in the secondary cluster are labeled.

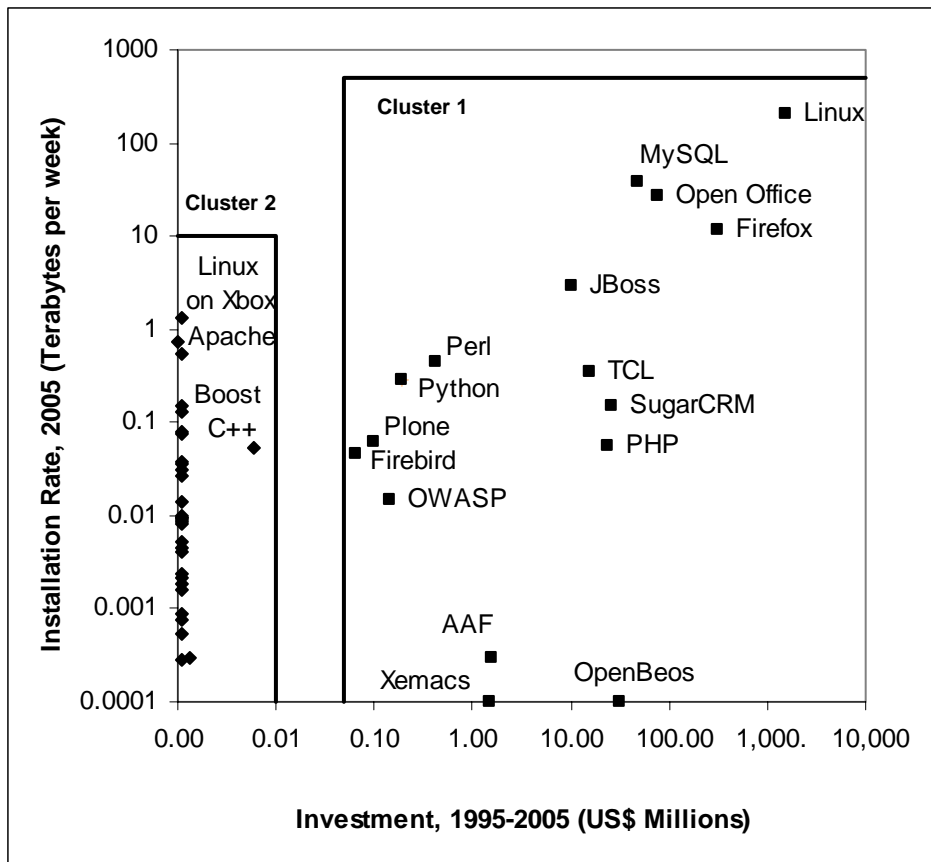**Figure 4: OSS Investment vs. Installation Rate[40]**

Figure 4 indicates the broader definition of the primary OSS cluster is more suitable, though NSLU, AAF, Xemacs, and OpenBeos require further examination due their significantly lower installation rates than the other projects in this cluster.

Figure 5 is a plot of the cumulative vendor investment in OSS projects by project for the projects in the sample set. Only every other project is labeled for clarity. The value for each project accounts for the investment in that project and all projects to the left. So, Linux accounts for 73% of vendor investments, Linux, Firefox (not labeled), and OpenOffice account for 92% of vendor investments, and so on… This plot shows how investments are heavily skewed to the top projects. Cluster 1 accounts for >99.99% of total investment and Cluster 2 accounts for <0.01% of total investment.

Figure 5: Cumulative OSS Investment[41]
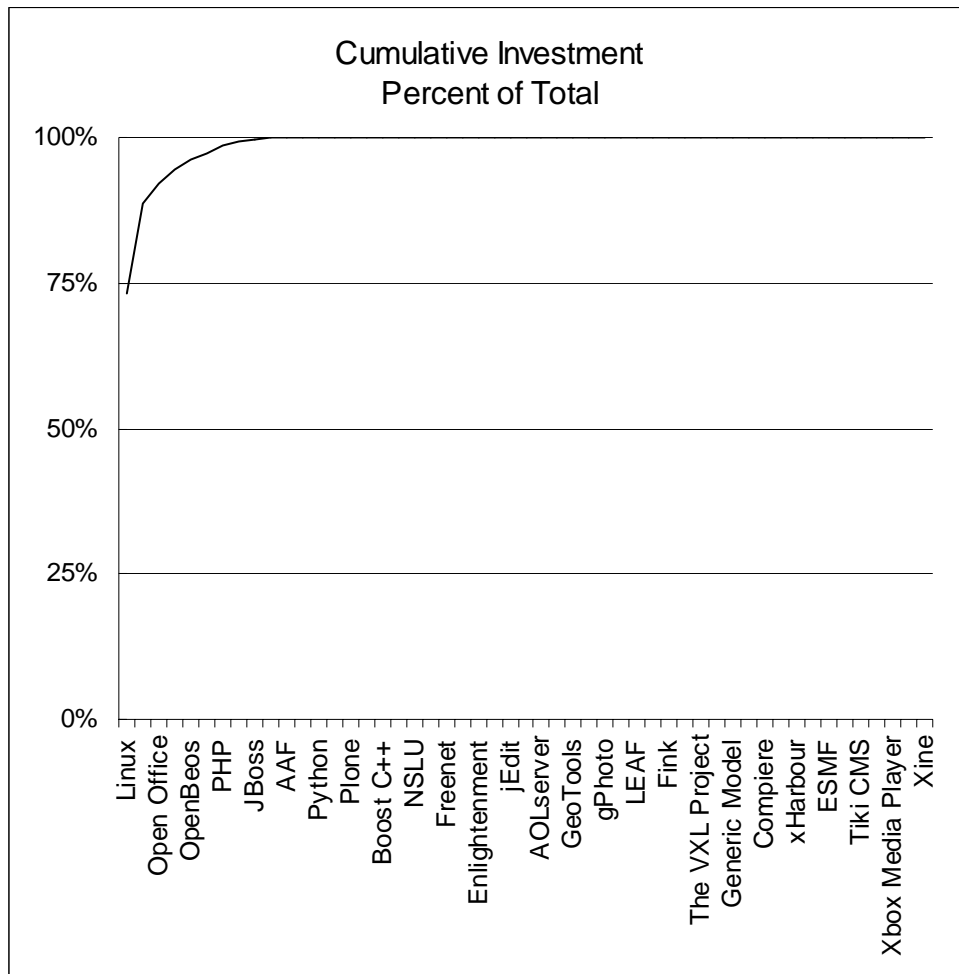
DISCUSSION

Cluster 1, the "money-driven cluster"

Cluster 1, consists of the eighteen projects that received more than 99% of vendor investments: Linux, Firefox, OpenOffice, MySQL, OpenBeos, SugarCRM, PHP, Tcl, JBoss, Xemacs, Advanced Authoring Format (AAF), Perl, Python, Open Web Application Security Project (OWASP), Plone, Freenet, NSLU, and Firebird . Excluding the two operating systems, Linux & OpenBeos, all projects run both Linux and Windows.

Four projects bear further scrutiny due to their significantly lower installation rates than the other projects in this cluster: NSLU, AAF, Xemacs, and OpenBeos. NSLU is software for the Linksys NSLU2 storage device, but is a community project that receives no support from Linksys[42]. AAF has received broad support from a variety of vendors, but has a low installation rate due to the small size of a file format and narrow target audience of manufacturers[43]. Xemacs investment has consisted of various companies paying a single software engineer to develop Xemacs as a contractor[44]. Finally, OpenBeos is a project that came out of a failed start-up[45].

Cluster 2, the "community-driven cluster"

Cluster 2, "the community-driven cluster", consists of 32 projects that almost no investment from IT Vendors. Most contributions would appear to be solely on a volunteer basis by their employees. Given the low levels of investment compare to Cluster 1, IT vendors' motivations are more likely to be altruistic than economic.

91% of these projects run on Linux and 75% run on Windows. Common wisdom would indicate that Linux would be the operating system of choice. Yet, there is significant usage of Windows, despite little public investment in OSS by Microsoft.

Motivations for Investment

If vendors were altruistically motivated to contribute to OSS projects, we would expect to see an even distribution of investments across projects. However, given the concentration of investment in Cluster 1, vendors more likely have economic motives for contributing to OSS projects. Wichmann identified four economic motivations: use OSS as a way to standardize IT components and move up the software stack to higher margin businesses, take advantage of OSS' price characteristics, use OSS to complement existing products, and enable compatibility to keep other products relevant.[46] Considering OSS business models[47,48] around distribution, service, and subscriptions, it seems most likely that vendors are investing in OSS as a complement to their existing IT product portfolios.
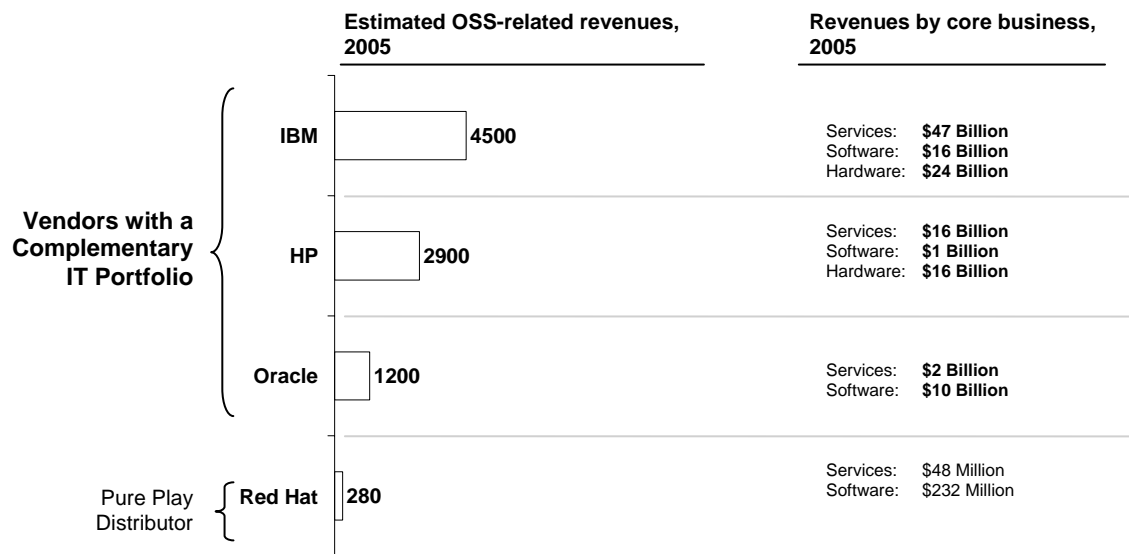
To test if vendors are using OSS as a complementary asset to drive revenues to their core businesses, we conduct two analyses. First, we examine how the amount of OSS-related revenues is related to the size of core businesses OSS can complement. Second, we look at IT vendors' actions towards OSS by software segment. If vendors are using OSS as a complement, then they likely avoid investing in OSS projects in the same software segments as their proprietary offerings. For these analyses, we examine HP, IBM, and Oracle: IT vendors with significant proprietary software revenues and significant OSS support. Red Hat is also examined as a contrasting example of a pure-play OSS vendor without non-OSS businesses that could be complemented by OSS.

IBM's and HP's OSS-related revenues were estimated by applying the Linux share of server revenues to their software and services businesses. IBM and HP's Linux server revenues are taken from IDC's quarterly server data. The revenues for Q1-Q2 2005[49] were used directly, and the revenues for Q3-Q4 2005 were estimated by applying

the year-on-year growth rate to the previous year's Linux revenues[50]. Oracle's OSS-related revenues were estimated applying the Linux share of database revenues as a percentage of software revenues to the company's total revenues. The breakdown of revenues between hardware, software, and services was derived from each vendor's annual report.[51]

Figure 6 is a plot of OSS-related revenues and core businesses' revenues for the IT vendors discussed above. The figure shows that vendors with the most OSS-related revenue also had the largest core businesses that could be complemented by OSS.

Figure 6: IT Vendor OSS-Related Revenues and Core Business Revenues

Estimated OSS-related revenues, 2005		Revenues by core business, 2005
IBM — 4500		Services: **$47 Billion** Software: **$16 Billion** Hardware: **$24 Billion**
HP — 2900		Services: **$16 Billion** Software: **$1 Billion** Hardware: **$16 Billion**
Oracle — 1200		Services: **$2 Billion** Software: **$10 Billion**
Red Hat — 280		Services: $48 Million Software: $232 Million

Vendors with a Complementary IT Portfolio: IBM, HP, Oracle

Pure Play Distributor: Red Hat

For the analysis of vendors' revenues and OSS activities by software segment, software segment revenues were generated from Gartner data and company annual reports. Revenues were grouped into five segments: Enterprise Applications, Middleware:Dev Tools, Middleware:Web and Application Server, Middleware:Database, and Operating Systems (OS). The Enterprise Applications segment includes Email & Calendaring, Business Intelligence, Content Management, Web Conferencing and Team

Collaboration, CRM, ERM, and CRM. The Middleware:Dev Tools segment includes

Application Development and Project and Portfolio Management. The Middleware:Web

and Application Server segment includes the Application Platform Suites, Application

Servers, B2B Middleware, Enterprise Service Bus, Integration Suite, Message Oriented

Middleware, Other AIM, and Portal Products markets. The Middleware:Database

segment includes both pre-relational and relational databases. HP was excluded from this

analysis due to unavailability of Gartner software segment data. Red Hat was also

excluded since it only produces open-source software, and there is no conflict with

existing software businesses.

  Vendors' actions were described as compete, collaborate, or contribute. Compete

is defined as actively promoting a proprietary product over an open source equivalent.

Collaborate is defined as helping distribute or market open source software, and possibly

offering some technical support, but not contributing code. Contribute is defined as

offering both support for OSS and source code.

  Figure 7 is a plot of IBM's revenues and OSS activities by software segment. It

shows the company's largest software segment by revenue is databases and they compete

against OSS equivalent by giving away a version of their proprietary software that has a

subset of the functionality available in the paid version. The bars indicating the size of

the revenue segments are shaded grey if IBM competes against OSS in that segment and

shaded with a diagonal cross-hatch if IBM both competes against and contributes to OSS

in that segment.

Figure 7: IBM's Revenues and OSS Strategies by Segment[52]

Software Revenues by Segment*
US$ Billions

Segment	Revenue	OSS Strategy	Details
Enterprise Applications	1.3	**Compete**	IBM has competing products: Lotus
Middleware: Dev Tools	1.8	**Compete/** Contribute	IBM has competing product: Rational, but also released Eclipse (Java application platform and IDE) into open source
Middleware: Web and Application Server	1.9	**Compete/** Contribute	IBM has competing product: Websphere. Source code is only available for light, Community version
Middleware: Database	4.1	**Compete**	Gives away a light version of DB2 (binary only)
OS	2.4	**Compete/** Contribute	IBM has competing product: AIX, but invests $100 MM yearly on Linux development, ported Jikes Compiler, Open Visualization, Data ExplorerVisualAge for Java, and MQseries to Linux

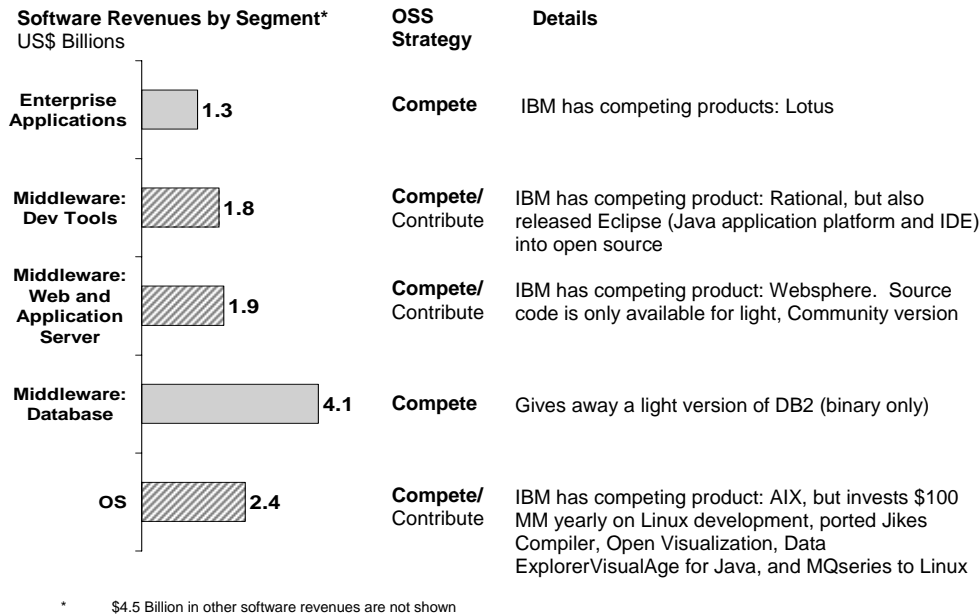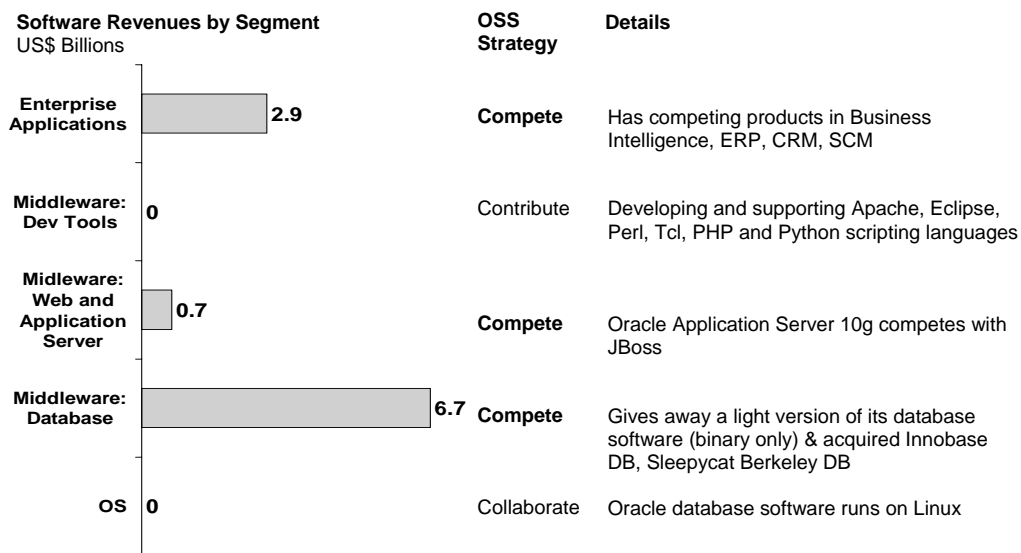* $4.5 Billion in other software revenues are not shown

Figure 8 is a plot of Oracle's revenues and OSS activities by software segment. It shows Oracle competes against OSS in software segments where they have revenues and either collaborate or contribute where they do not. The bars indicating the size of the revenue segments are shaded grey if Oracle competes against OSS in that segment.

Figure 8: Oracle's Revenues and OSS Strategies by Segment[53]

Software Revenues by Segment
US$ Billions

Segment	Revenue	OSS Strategy	Details
Enterprise Applications	2.9	**Compete**	Has competing products in Business Intelligence, ERP, CRM, SCM
Middleware: Dev Tools	0	Contribute	Developing and supporting Apache, Eclipse, Perl, Tcl, PHP and Python scripting languages
Midlware: Web and Application Server	0.7	**Compete**	Oracle Application Server 10g competes with JBoss
Middleware: Database	6.7	**Compete**	Gives away a light version of its database software (binary only) & acquired Innobase DB, Sleepycat Berkeley DB
OS	0	Collaborate	Oracle database software runs on Linux

For each vendor, a software segment is considered a core business if it is in the top three, when ranked by revenues, and a non-core business if it is in the bottom two.

Table 3: IBM and Oracle's Core Segments and OSS Strategies

Vendor	Core software segments	Does vendor compete with open source software in that segment?
IBM	Middleware: Database	Yes
	Operating Systems	Yes
	Middleware: Web and Application Server	Yes
Oracle	Middleware: Database	Yes
	Enterprise Applications	Yes
	Middleware: Web and Application Server	Yes

Table 3 is a summary of each vendor's OSS activities for their core software segments. It shows IBM and Oracle compete with OSS software in 6 of 6 core software segments.

Table 4: IBM and Oracle's Non-Core Segments and OSS Strategies

Vendor	Non-Core software segments	Does vendor compete with open source software in that segment?
IBM	Enterprise Applications	Yes
	Middleware: Development Tools	Yes
Oracle	Middleware: Development Tools	No
	Operating Systems	No

Table 4 is a summary of each vendor's OSS activities for their non-core software segments. It shows IBM and Oracle only compete with open source software in 2 of 4 non-core software segments.

Tables 3-4 show vendors are more likely to compete with OSS in core segments than in non-core software segments, consistent with the hypothesis that vendors are using OSS as an asset complementary to their core businesses. However, Figure 7 indicates vendors sometimes also invest in OSS in a core software segment, as IBM has done with Linux in Operating Systems. That investment seems to create a conflict with IBM's proprietary operating system, AIX, in contradiction to the hypothesis that vendors are using OSS as a complementary asset. However, company statements and news articles[54] reveal IBM is still using OSS (Linux) as a complement to its core business (AIX), because it is positioning AIX as the easiest and most compatible upgrade path for Linux:

- "IBM is investing equally in Linux and AIX and in making the two operating systems work well together." - Rich Michos, VP Linux Servers, IBM

- "The next version of IBM's AIX—long regarded as one of the most proprietary of the commercial Unices--goes the farthest in melding Linux technologies with an older, established version of UNIX." – Linux Planet.com

- "AIX 5L is really designed to extend Linux past its 32-bit roots and give enterprises an upgrade path." - Andy Wachs, Unix Strategy Director, IBM

We see that vendors can not only invest in OSS projects that immediately complement their IT portfolios but also modify their portfolios so the OSS project becomes complementary.

Conclusions

The most important findings in this study are the grouping of OSS projects into clusters and the identification of IT vendors' motives in each cluster. Individuals' motivations for contributing to OSS had been found to be a mix of intrinsic and altruistic.[55] For vendors, the equivalent of an intrinsic motivation is an altruistic motivation, and the equivalent of an extrinsic motivation is an economic motivation. The grouping of projects identified a "money-driven cluster" where IT vendors' motives are economic. In this cluster, significant investments have been made in projects that will serve as complementary assets to drive revenues to vendors' core businesses. In the "community-driven cluster", monetary investment from IT vendors is limited, and IT vendors appear to be altruistically motivated in their contributions.

This study also proposed and tested a new metric for the impact of an OSS project: installation rate, which is defined the number of weekly distributions times the size of the installation file. This metric was one of the variables used to cluster projects and is an improvement over previously proposed OSS metrics. It is quantitative, objective, and applicable across OSS projects. However, it may give a skewed picture for segments which require frequent updates (e.g., file-sharing, security).

The implications of these findings merit further investigation. The projects in "money-driven cluster" had greater impact, as measured by installation rate, and received greater investment. Should the correlation between impact and investment change how one thinks about "free" software? How does the impact of open source software change with investment? If companies are extrinsically motivated, how will the participation and investment of IT vendors in OSS affect the quality of the OSS and utility for the

community in the long-term? Does "code written for private needs serves public needs less well"?[56] If vendors are using a complementary asset business model, will development become focused less on new software functionality and more on hardware and services (e.g., integration tools, integration to proprietary hardware)? In both clusters, Windows and Linux are supported to almost equal degrees. Does the prevalence of Windows in OSS projects indicate the platform of choice for OSS is an open question?

REFERENCES

[1] Paper number is a placeholder until actual number is assigned by HBS Working Paper Publishing.

[2] Crowston, K., Annabi. H, and Howison, J., "Defining Open Source Software Project Success", 2003, Syracuse School of Information Studies Working Paper

[3] Harhoff, D., Henkel, J., and E. von Hippel. "Motivations for participating in open-source projects," *International Journal of Electronic Commerce*, Vol 6 No 3, 2003, pp. 25-39

[4] Nakakoji, K., Yamamoto, Y., Nishinaka, Y., Kishida, K., Ye, Y. "Evolution Patterns of Open-Source Software Systems and Communities," in Proceedings of the International Workshop on Principles of Software Evolution, Session 4, 2002, pp.76-85

[5] Rossi, C. and Bonaccorsi, A, "Contributing to the common pool resources in Open Source software. A comparison between individuals and firms," 2004, Working Paper, Sant'Anna School of Advanced Studies Institute for Informatics and Telematics (IIT-CNR), Pisa, Italy

[6] Bates, J., Di Bona, C., Lakhani, K., Wolf B. The Boston Consulting Group hacker survey, 2002 (available online at http://www.osdn.com/bcg/BCGHACKERSURVEY-0.73.pdf)

[7] Bitzer, J., Schrettl, W., Schroder, P. "Intrinsic Motivation in Open Source Software Development," 2004. (available online at: http://129.3.20.41/eps/dev/papers/0505/0505007.pdf)

[8] Ghosh, R.A., and Prakash, V.V. "The Orbiten Free Software Survey: May 2000," 2002a, First Monday (Peer-reviewed journal on the internet) http://firstmonday.org/issues/issue5_7/ghosh/

[9] Ghosh R.A., Glot, R., Krieger, B., Robles, G. "Survey of Developers," in *Free/Libre and Open Source Software: Survey and Study, FLOSS Final Report*, International Institute of Infonomics, Berlecom Research GmbH, June 2002b. (available online at: http://www.infonomics.nl/FLOSS/report/)

[10] Lakhani, K.R., and Wolf, R.G. "Why hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects," in *Perspectives on Free and Open Source Software*, J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani (eds.), MIT Press, Cambridge, MA, 2005, pp. 3-21

[11] Wilcox, J. "IBM to spend $1 billion on Linux in 2001", CNET news.com, December 12, 2000. (available online at: http://news.com.com/2100-1001-249750.html)

[12] Lerner, J., and Tirole, J. "Economic Perspectives on Open Source," in *Perspectives on Free and Open Source Software*, J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani (eds.), MIT Press, Cambridge, MA, 2005, pp. 47-78

[13] Rossi, C. and Bonaccorsi, A. "Why profit-oriented companies enter the OS field? Intrinsic vs. extrinsic incentives," paper presented for Fifth Workshop on Open source Software Engineering (5-WOSSE) May 17, 2005, St Louis, MO, USA.

[14] Deci, E., and Ryan, R. Intrinsic Motivation and Self-Determination in Human Behavior, Plenum Press, New York, NY, 1985

[15] Ryan, R.M., and Deci, E.L. "Intrinsic and extrinsic motivations: Classic definitions and new directions," *Contemporary Educational Psychology*, Vol 25, 2000, pp.54-67

[16] Frey, B. *Not Just for the Money: An economic theory of personal motivation*. Edward Elgar, Brookfield, VT, 1997

[17] Lerner, J., and Tirole, J. "Some Simple Economics of Open Source," *The Journal of Industrial Economics,* Vol 50, No 2, June 2002, pp. 197-234

[18] Lakhani, K.R., and Wolf, R.G. "Why hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects," in *Perspectives on Free and Open Source Software*, J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani (eds.), MIT Press, Cambridge, MA, 2005, pp. 3-21

[19] Lerner, J., and Tirole, J. "Some Simple Economics of Open Source," *The Journal of Industrial Economics,* Vol 50, No 2, June 2002, pp. 197-234

[20] Holmstrom, B., 1999, "Managerial Incentive Problems: A Dynamic Perspective" *Review of Economic Studies* 66 (January), pp. 169-182

[21] Ghosh R.A., Glot, R., Krieger, B., Robles, G. "Survey of Developers," in *Free/Libre and Open Source Software: Survey and Study, FLOSS Final Report*, International Institute of Infonomics, Berlecom Research GmbH, June 2002b

[22] Rossi, C. and Bonaccorsi, A. "Why profit-oriented companies enter the OS field? Intrinsic vs. extrinsic incentives," paper presented for Fifth Workshop on Open source Software Engineering (5-WOSSE) May 17, 2005, St Louis, MO, USA.

[23] Wichmann, T. "Basics of Open Source software markets and business models," in *Free/Libre and Open Source Software: Survey and Study, FLOSS Final Report*, International Institute of Infonomics, Berlecom Research GmbH, June 2002. (available online at: http://www.infonomics.nl/FLOSS/report/)

[24] Rossi, C. and Bonaccorsi, A. "Why profit-oriented companies enter the OS field? Intrinsic vs. extrinsic incentives," paper presented for Fifth Workshop on Open source Software Engineering (5-WOSSE) May 17, 2005, St Louis, MO, USA.

[25] Nakakoji, K., Yamamoto, Y., Nishinaka, Y., Kishida, K., Ye, Y. "Evolution Patterns of Open-Source Software Systems and Communities," in Proceedings of the International Workshop on Principles of Software Evolution, Session 4, 2002, pp.76-85

[26] Crowston, K., Annabi. H, and Howison, J., "Defining Open Source Software Project Success", 2003, Syracuse School of Information Studies Working Paper

[27] Source: www.wikipedia.org

[28] Wichmann, T. "Basics of Open Source software markets and business models," in *Free/Libre and Open Source Software: Survey and Study, FLOSS Final Report*, International Institute of Infonomics, Berlecom Research GmbH, June 2002.

[29] DeLone, W. H., and McLean, E. R., 1992, "Information Systems Success: The Quest for the Dependent Variable", *Information Systems Research, 3(1)*, pp. 60-95

[30] Crowston, K., Annabi. H, and Howison, J., "Defining Open Source Software Project Success", 2003, Syracuse School of Information Studies Working Paper

[31] Ibid and references within

[32] Multiple reports by IDC and Gartner. Specific reports are referenced later in this paper.

[33] IDC, "The Linux Marketplace – Moving From Niche to Mainstream", Dec 14, 2004 (available at http://www.osdl.org/docs/linux_market_overview.pdf)

[34] StatSoft, Inc. (2006). Electronic Statistics Textbook. Tulsa, OK: StatSoft. (available online at: http://www.statsoft.com/textbook/stathome.html)

[35] Source: Wikipedia

36 Source: IDC, "The Linux Marketplace – Moving From Niche to Mainstream"; www.netcraft.com; www.internetnews.com; www.news.com; www.jboss.com; www.itjungle.com; www.ddj.com; www.mozilla.org; www.opensource.org; www.perl.org; www.newsforge.com; www.python.org; www.sourceforge.net; www.codeplex.com; http://www.consortiuminfo.org/standardsblog/article.php?story=20060513194119192; http://www.mysql.com/news-and-events/on-demand-webinars/scaleout-2006-05-09.pdf; http://perl.apache.org/outstanding/stats/netcraft.html;

[37] Valued at $150,000 per year

[38] Source: Wichmann (2002); http://www.mozilla.org/press/mozilla-foundation.html; Mozilla Foundation IRS tax form 990, 2003-2004; Split between Netscape client and server development was based on respective percentage of revenues as estimated from Hambrecht and Quist research notes,1995-1998; Sun 10-Q filing, Q1, 2000; Sun Annual Report, 2000; Sun 10-K filing, 2001; http://www.sun.com/smi/Press/sunflash/2000-10/sunflash.20001016.4.xml; http://www.mysql.com/news-and-events/press-release/release_2006_10.html; http://www.mysql.com/news-and-events/press-release/release_2003_18.html; http://www.mysql.com/news-and-events/press-release/release_2001_3.html; http://www.techworld.com/news/index.cfm?newsID=3510; http://www.sugarcrm.com/crm/about/press-releases/20040802-funding.html; http://www.sugarcrm.com/crm/about/press-releases/20050215-sugarcrm-series-b.html; http://www.sugarcrm.com/crm/about/press-releases/20051012-funding.html; http://news.com.com/JBoss+lands+10+million+in+funding/2100-1014_3-5161753.html; www.opensource.org; www.perl.com; perl.apache.org; http://httpd.apache.org/ABOUT_APACHE.html; www.newsforge.com; www.redherring.com; www.python.org Interwoven annual report, 2005; Interwoven 8-K filing, Oct. 31, 2000; PENDER Financial press release, Feb. 21, 2006; www.enlightenment.com; www.xemacs.org; http://www.crystalspace3d.org/tikiwiki/tiki-index.php?page=Project+Donations; http://freenetproject.org/news.html; http://www.squid-cache.org/SPONSORS.txt/; private communication with Dave Abrahams for Boost C++; https://www.osgeo.org/;

http://www.firebirdsql.org/index.php?op=ffoundation&id=sponsorship;
http://www.aafassociation.org/html/Memberlist.htm; http://www.aafassociation.org/html/join.html;
http://moodle.org/donations/; http://www.beatjapan.org/mirror/www.be.com/aboutbe/pressreleases/98-11-16_befunding.html; http://plone.org/foundation/faq; private communication with Marc Laporte for Tiki CMS; http://old.owasp.org/about/membership/corporate.html;
http://xboxmediacenter.com/info_contrib.htm; http://www.nslu2-linux.org/wiki/Info/DonationList;
www.guidestar.org

[39] Source: Cluster and TreeView software, M. Eisen, available at rana.stanford.edu/software

[40] Ibid

[41] Ibid

[42] Source: http://www.nslu2-linux.org/wiki/Info/DonationList

[43] Source: http://www.aafassociation.org/html/faq.html

[44] Source: http://www.xemacs.org/People/ben.wing/

[45] Source: http://haiku-os.org; http://www.beatjapan.org/mirror/www.be.com/aboutbe/pressreleases/96-04-15_Financing.html; http://www.beatjapan.org/mirror/www.be.com/aboutbe/pressreleases/98-11-16_befunding.html

[46] Wichmann, T. "Basics of Open Source software markets and business models," in *Free/Libre and Open Source Software: Survey and Study, FLOSS Final Report*, International Institute of Infonomics, Berlecom Research GmbH, June 2002

[47] Koenig, J., "Seven open source business strategies for competitive advantage", in IT Manager's Journal, May 14, 2004. http://management.itmanagersjournal.com/management/04/05/10/2052216.shtml?tid=85

[48] Krishnamurthy, S., "An Analysis of Open Source Business Models," in *Perspectives on Free and Open Source Software*, J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani (eds.), MIT Press, Cambridge, MA, 2005, pp. 279-296

[49] Source: IDC Press Release, "Worldwide Server Market Shows Growing IT Investment Across Platforms", August 26, 2005 (available at http://www.idc.com/getdoc.jsp?containerId=prUS00223005); IDC Press Release, "Worldwide Server Market Continues Growth in First Quarter as Windows and Unix Server Markets Generate Equal Revenues for First Time", May 26, 2005 (available at http://www.idc.com/getdoc.jsp?containerId=prUS00153905); ZDNet, "IBM King of Server Revenue", August 25, 2005 (available at http://blogs.zdnet.com/BTL/?p=1780)

[50] Source: Network World, "Linux server sales top $1 billion in Q3", Nov 24, 2004 (available at http://www.networkworld.com/news/2004/1124linuxserve.html); IDC Press Release, "Worldwide Server Market Slows In Fourth Quarter But Grows To $51.3 Billion In 2005" February 22, 2006 (available at http://www.idc.com/getdoc.jsp?containerId=prUS20074406); IDC Press Release, "Worldwide Server Market Shows Strongest Growth in the Volume and Midrange Server Segments", November 23, 2005 (available at http://www.idc.com/getdoc.jsp?containerId=prUS20013705); IT Jungle, "IDC Says Linux Server Market Grew 36 Percent In Q4 2004", March 8, 2005 (available at http://www.itjungle.com/tlb/tlb030805-story01.html)

[51] Source: Oracle Press Release, "Oracle is #1 on Linux" (available at http://www.oracle.com/database/feature_db_dbleadership.html); IBM, HP, Oracle, Red Hat annual reports, 2005

[52] Source: IBM annual report, 2005; www.eclipse.org; Gartner "Market Share: Enterprise Email and Calendaring, Worldwide, 2003-2005", June 2006; Gartner "Market Share: Business Intelligence Platforms, Worldwide, 2005", September 2006; Gartner "Market Share: Enterprise Content Management Software, Worldwide, 2003-2005", May 2006; Gartner "Market Share: Web Conferencing and Team Collaboration, Worldwide 2003-2005", August 2006; Gartner "Market Share: CRM Software, Worldwide, 2005", June 2006; Gartner "Market Share: ERP Software, Worldwide, 2005", June 2006; Gartner "Market Share: Supply Chain Management Software, Worldwide, 2005", June 2006; Gartner "Market Share: Application Development and Project and Portfolio Management, Worldwide, 2005, Update", September 2006; Gartner "Market Share: AIM and Portal Software, Worldwide, 2005", June 2006; Gartner "Market Share: Database Management Systems Software, Worldwide, 2005", June 2006

[53] Source: Oracle annual report, 2005; Oracle press releases; SearchOpenSource, Nov. 3, 2005; Gartner "Market Share: Enterprise Email and Calendaring, Worldwide, 2003-2005", June 2006; Gartner "Market Share: Business Intelligence Platforms, Worldwide, 2005", September 2006; Gartner "Market Share:

Enterprise Content Management Software, Worldwide, 2003-2005", May 2006; Gartner "Market Share: Web Conferencing and Team Collaboration, Worldwide 2003-2005", August 2006; Gartner "Market Share: CRM Software, Worldwide, 2005", June 2006; Gartner "Market Share: ERP Software, Worldwide, 2005", June 2006; Gartner "Market Share: Supply Chain Management Software, Worldwide, 2005", June 2006; Gartner "Market Share: Application Development and Project and Portfolio Management, Worldwide, 2005, Update", September 2006; Gartner "Market Share: AIM and Portal Software, Worldwide, 2005", June 2006; Gartner "Market Share: Database Management Systems Software, Worldwide, 2005", June 2006

[54] Source: Eweek.com; LinuxPlanet.com; Freeos.com, LinuxPlanet.com

[55] Ghosh R.A., Glot, R., Krieger, B., Robles, G. "Survey of Developers," in *Free/Libre and Open Source Software: Survey and Study, FLOSS Final Report*, International Institute of Infonomics, Berlecom Research GmbH, June 2002b

[56] Harhoff, D., Henkel, J., and E. von Hippel. "Motivations for participating in open-source projects," *International Journal of Electronic Commerce*, Vol 6 No 3, 2003, pp. 25-39