

Collaborative Software Design & Development

Comparing Free Open Source and Proprietary Products

Enos Jones and Danhua Shao

Overview

- Background on innovation and invention in software
- Review of literature
- Data and Methodology
- Empirical results
- Critique

Top 10 software innovations of 2004

- 10. Connect 2nd monitor via the Network
- 9. Map24
- 8. Open Office 1.1 - Print to PDF, Flash
- 7. CodeRush for Visual Studio .NET
- 6. Visual Studio 2005/C#/VB.NET/IDE/Framework 2.0
- 5. Gmail - Client side web app.
- 4. Keyhole - Satellite imagery
- 3. Google Desktop
- 2. Microsoft Office System 2003 - Research Sidebar/Pane
- 1 software innovation - Mozilla Firefox - Find as you type.

- Source: Robert Gelb <http://www.vbrad.com/article.aspx?id=64>

Research Paper

→ Authors

↳ Dario Lorenzi - Professor

↳ Cristina Rossi - Researcher



→ Department of Management, Economics, and Industrial Engineering, Politecnico di Milan

What is Innovation in software?

- Widely Debated by scholars
- What is the boundary line between innovation and not in programs?
- Is adaptation to different platforms innovation?
- What are the most important elements?
- Seminal studies have highlighted distinction between invention and innovation
 - ↳ Invention: first occurrence of idea for new product or process
 - ↳ Innovation: its first commercialization
- Innovation - least analyzed economic part despite contributing most to economic growth

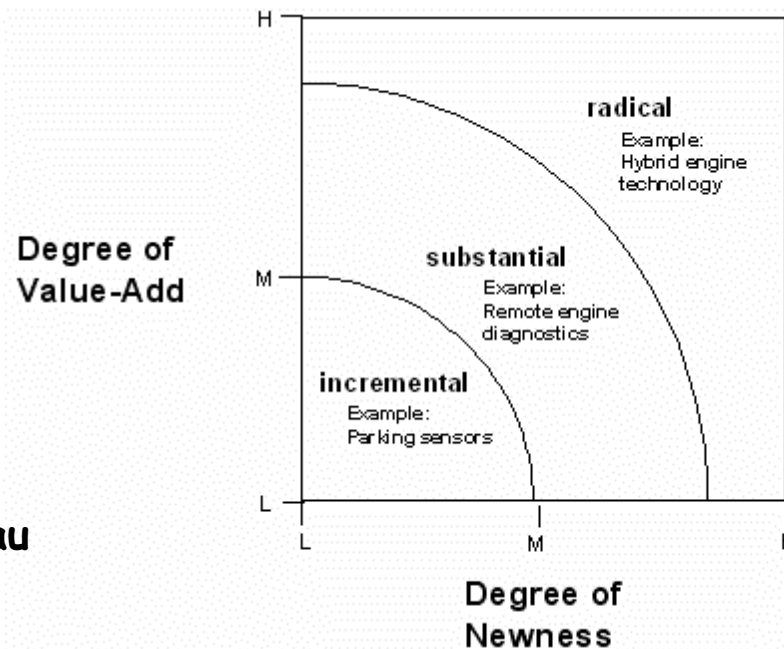
Classifying Innovation

→ Five different classes

↳ New products; new methods of production; new sources of supply; exploitation of new markets; new ways to organize business

→ Incremental and radical innovation

→ Three levels



Paul Wright - Invetech
<http://www.ceoforum.com.au>

Classifying Innovation

→ No comprehensive indicator

→ Metrics

↳ R&D activity

↳ Patents

↳ Bibliometrical data(publications)

→ Collect data through surveys and databases

Failures of Old Indicators

- the New Economy has made traditional measures of innovation obsolete
 - ↳ Patents
 - ↳ Trademarks
- From Commodity-driven to knowledge-driven economy creates new variables
 - ↳ Creation of knowledge
 - ↳ Management of knowledge
 - ↳ Transfer of knowledge through society
- Patents
 - ↳ Can't keep up with rapid evolution
 - ↳ Software companies use alternative protection
 - ↳ Process is in need of change

Need for New Indicators

→ Account for general software aspects

- ↳ Drivers of internal innovation process
- ↳ Comparisons with other market solutions

→ Capture specific elements

- ↳ Programming languages
- ↳ Platforms
- ↳ Software architecture
- ↳ New coding
- ↳ Library reuse

Classification of Innovation for FOSS

- Radical innovation
 - Technological modifications
 - Platform modifications
 - Market innovations
-
- Based on this classification taxonomy analyzed 500 SourceForge projects
 - ↳ 436 - not innovative
 - ↳ 5 - radical innovation

Data and Methodology

→ **Sample from study of >900 software firms**

↳ Finland, Germany, Italy, Portugal and Spain

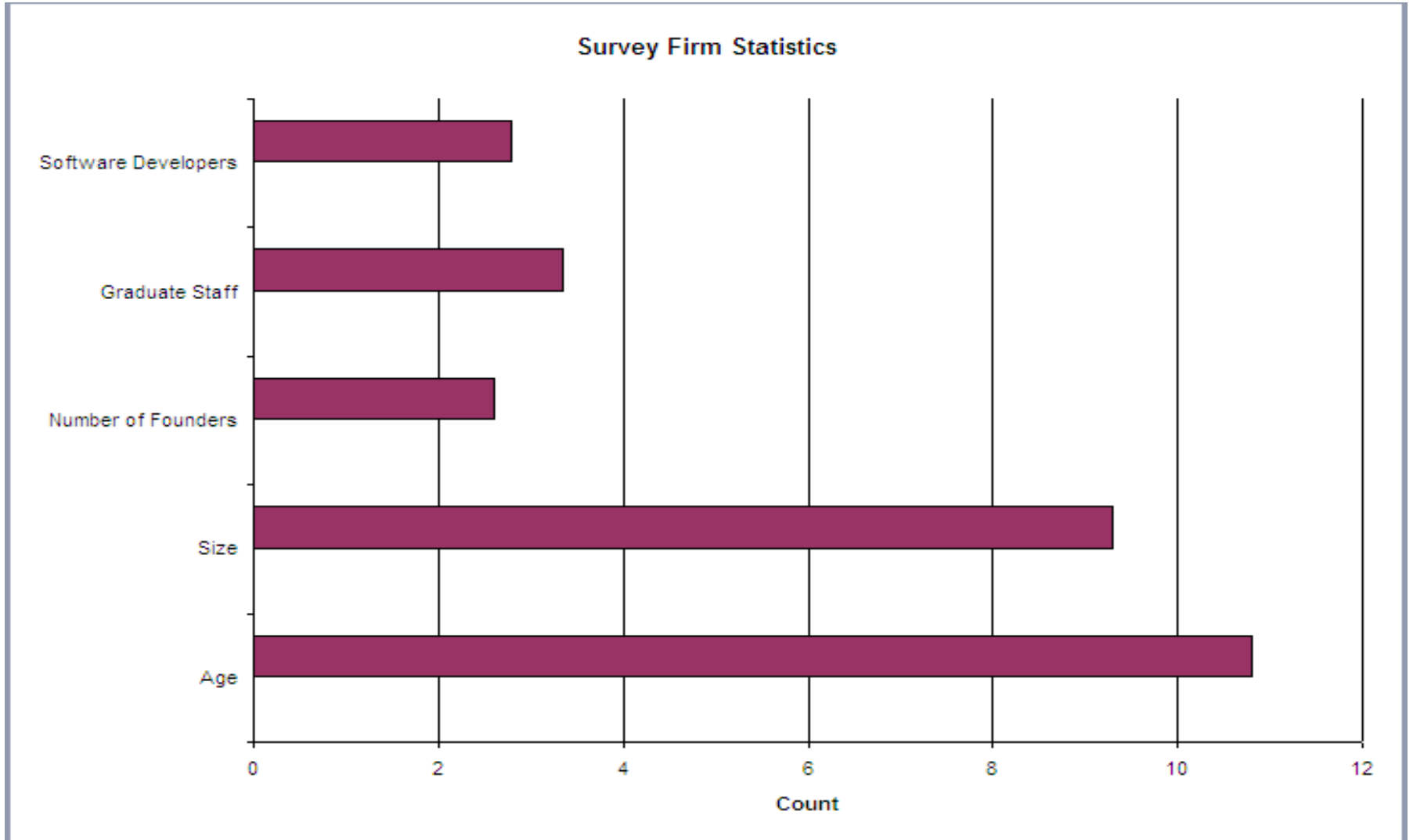
↳ Large survey - 2004

→ **Focused on Italian Respondents**

↳ Available database links complete for Italy only

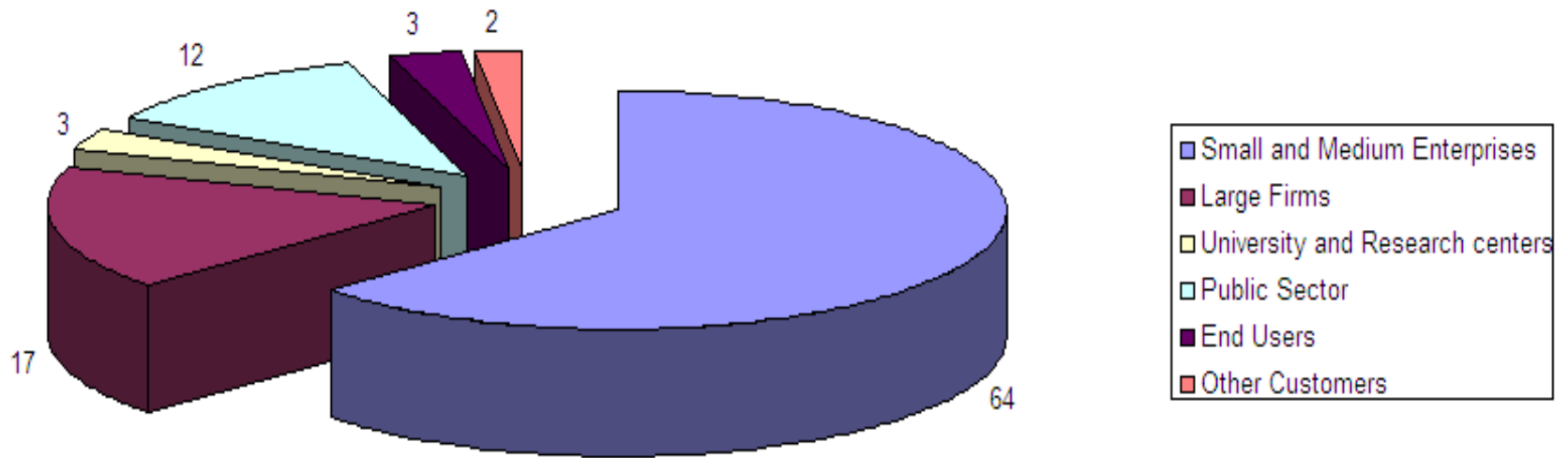
↳ Explore Small and Medium Enterprises with low sales and R&D

Characteristics of Italian Firms



Customers of Firms

Firms' Customers



Assessment Method 1

→ Gathered Data from 3 traditional sources:

→ Trademarks

↳ Database of the European Office for Trademarks and Design

→ Patents

↳ Delphion database

→ Scientific and Technical Publications

↳ Scopus

→ As a means of verifying need for alternative indicators

Assessment Method 2

→ Evaluation by a group of experts (3)

↳ Graduate in Telecommunications Engineering working in large firm

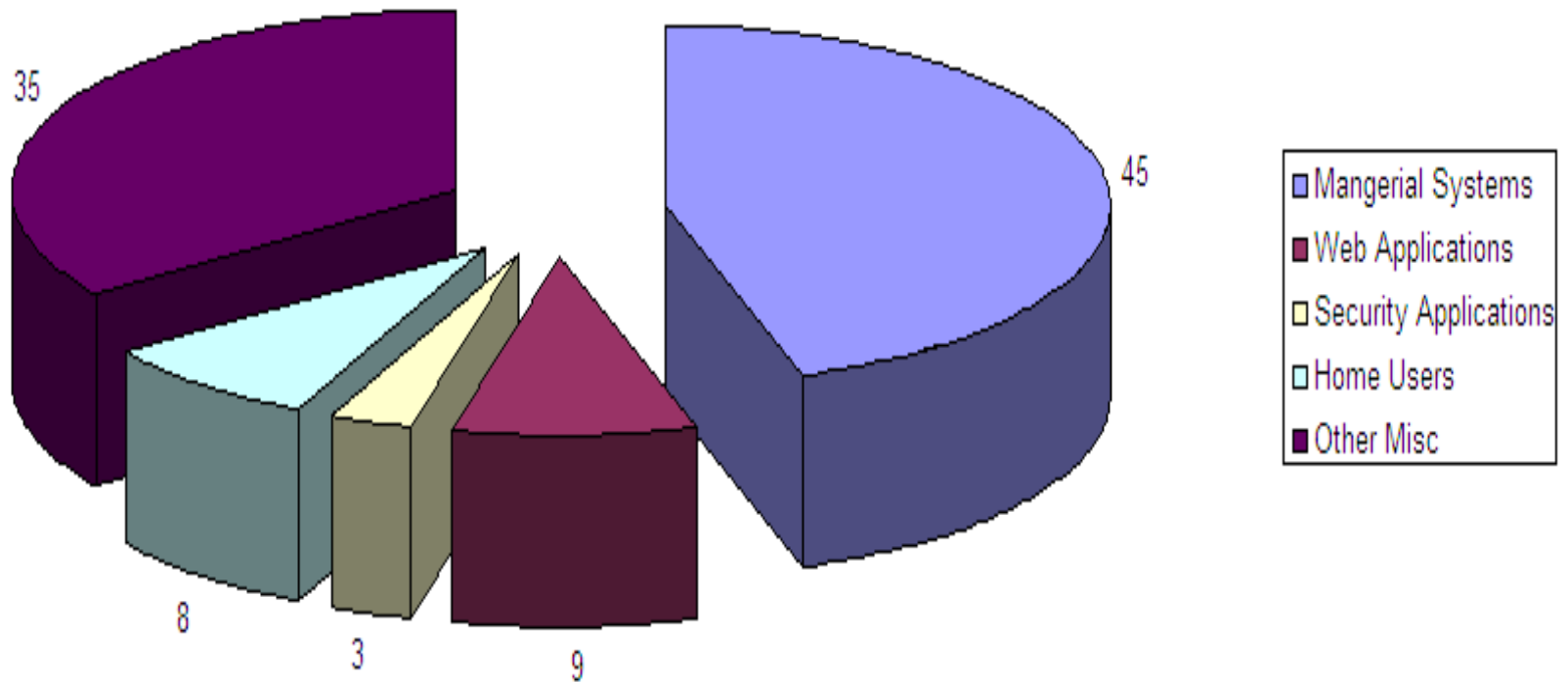
↳ Graduate in Computer Science, working as System Administrator

↳ Graduate in Computer Science, employed as Research Assistant

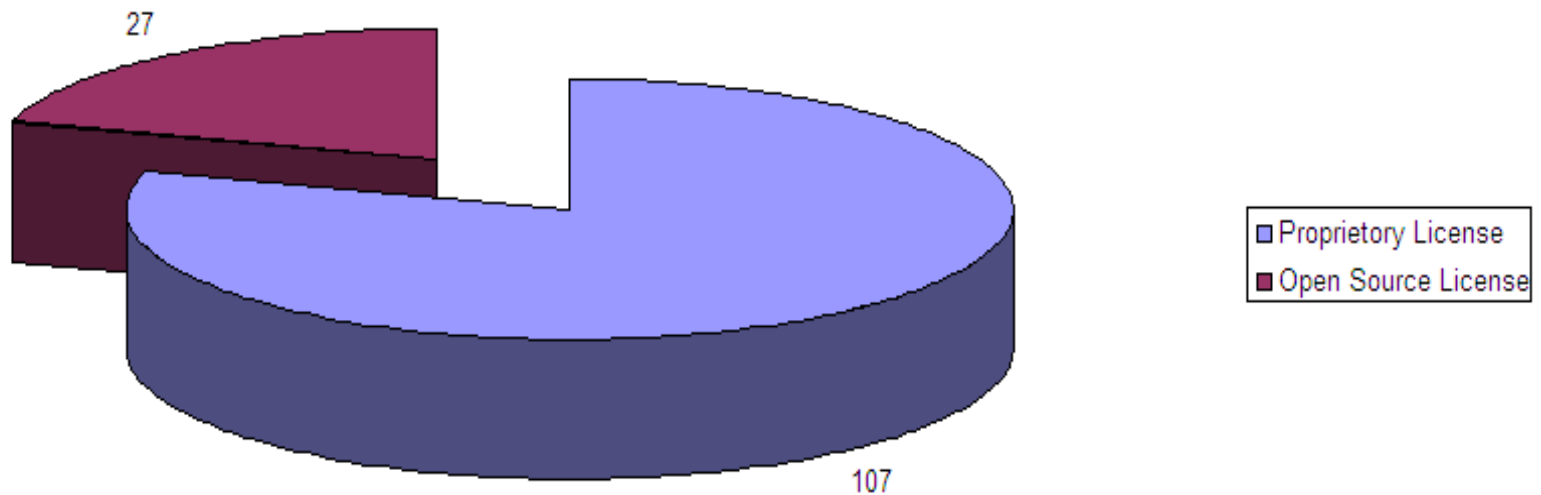
→ Each asked to evaluate innovativeness of 134 software solutions

→ Each after in depth introduction used a table to reflect opinions

Software Solution Characteristics



Types of Software License in Sample



Experts' Table

Table 3: Example of a record of experts' table

<i>Field</i>	<i>Content</i>
Product index	081
Firm index	051
Name of the product (with link to its Web page)	XXX
Product typology	Software for e-commerce
Brief description of product	Web oriented application for the business to business commerce, targeted to the fashion sector; possibility of acquisition of orders and real time checks; interoperability among different databases (postgresql, ms-SQL, IBM DB2, Oracle) and operative systems (it runs on both Windows and Linux); multi-language; modular structure; possibility of documents tracking; based on the managerial system as400 or stealth v.3; it uses a php 4 technology; possibility of SSL cryptography
Brief description of firm	XXX: development of Web oriented applications, in particular targeted to the fashion sector; development of software for e-learning and content management
Indicator 1: product new to the firm	2
Eventual notes on Indicator 1	-----
Indicator 2: product new to the market in what it does	2
Eventual notes on Indicator 2	-----
Indicator 3: product new to the market under technological viewpoint	4
Eventual notes on Indicator 3	-----
Indicator 4: modules new to the world	2
Eventual notes on Indicator 4	-----
Indicator 5: platform new to the world	3
Eventual notes on Indicator 5	-----

Indicators

→ First Dimension (internal level)

↳ Indicator 1: Product New To Firm?

- Compare software with others developed by same firm
- Highlight peculiarities and differences

→ Second Dimension(firm's market)

↳ Indicator 2: Product New to Market in what it does?

- Compare solution with others in market
- Innovation in what a solutions does

↳ Indicator 3: Is the Product new to the market under a technological viewpoint?

- Innovation in how a solution accomplishes task
- Exploration of technical aspects

→ Third Dimension(world level)

↳ Indicator 4: Is the software innovative as it contains peculiar and original modules, which can be hardly found in other solutions?

Indicators (cont.)

→ Third Dimension(world level)

↳ Indicator 5: How the software platform is new with respect to those of other software solutions, also in different market segments?

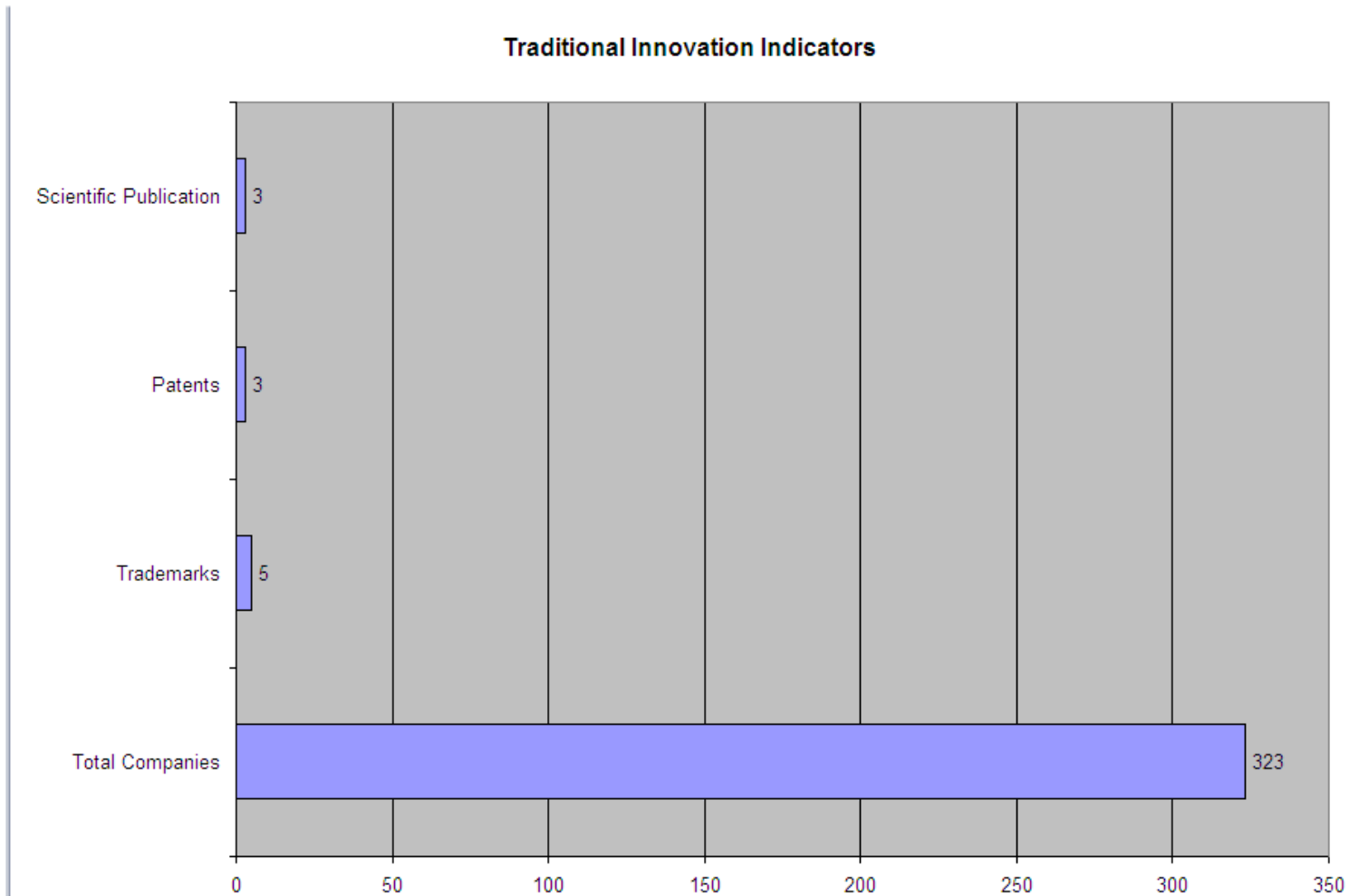
→ Shortcomings

- ↳ Subjective nature of Experts' opinions
- ↳ Size of group of experts

→ Advantages

- ↳ Experts' in-depth market and multiple opinions
- ↳ Statistical correlation indicated consistency for experts' metrics
- ↳ Avoided problems with Surveys having respondents whose knowledge is insufficient to task
- ↳ Summed experts opinions on each indicator to preserve variability and reducing outliers

Empirical Results



Inference from Traditional Results

→ Traditional methods not suitable to capture innovation

- ↳ Low use of patents
- ↳ Low use of trademarks
- ↳ Few publications

→ This can be a means of justifying an alternative methodology

- ↳ Lack of information from traditional indicators on innovations

Empirical Results - "Experts" Method

→ Measurements through research and assessment

↳ Innovation within firm (Dimension 1)

➤ Is the product new to the firm (Indicator 1)

↳ Market in which firm operates (Dimension 2)

➤ Does software better satisfy user needs? (Indicator 2)

➤ Is this new technology to the market (Indicator 3)

↳ Generality of software innovation to state-of-art and knowledge (Dimension 3)

➤ Does it contain peculiar and original modules?(Indicator 4)

➤ Newness of software platform to other solutions (Indicator 5)

Grading System for Experts

→ Range of scores

↳ 1 (not at all innovative)

↳ 5 (very innovative)

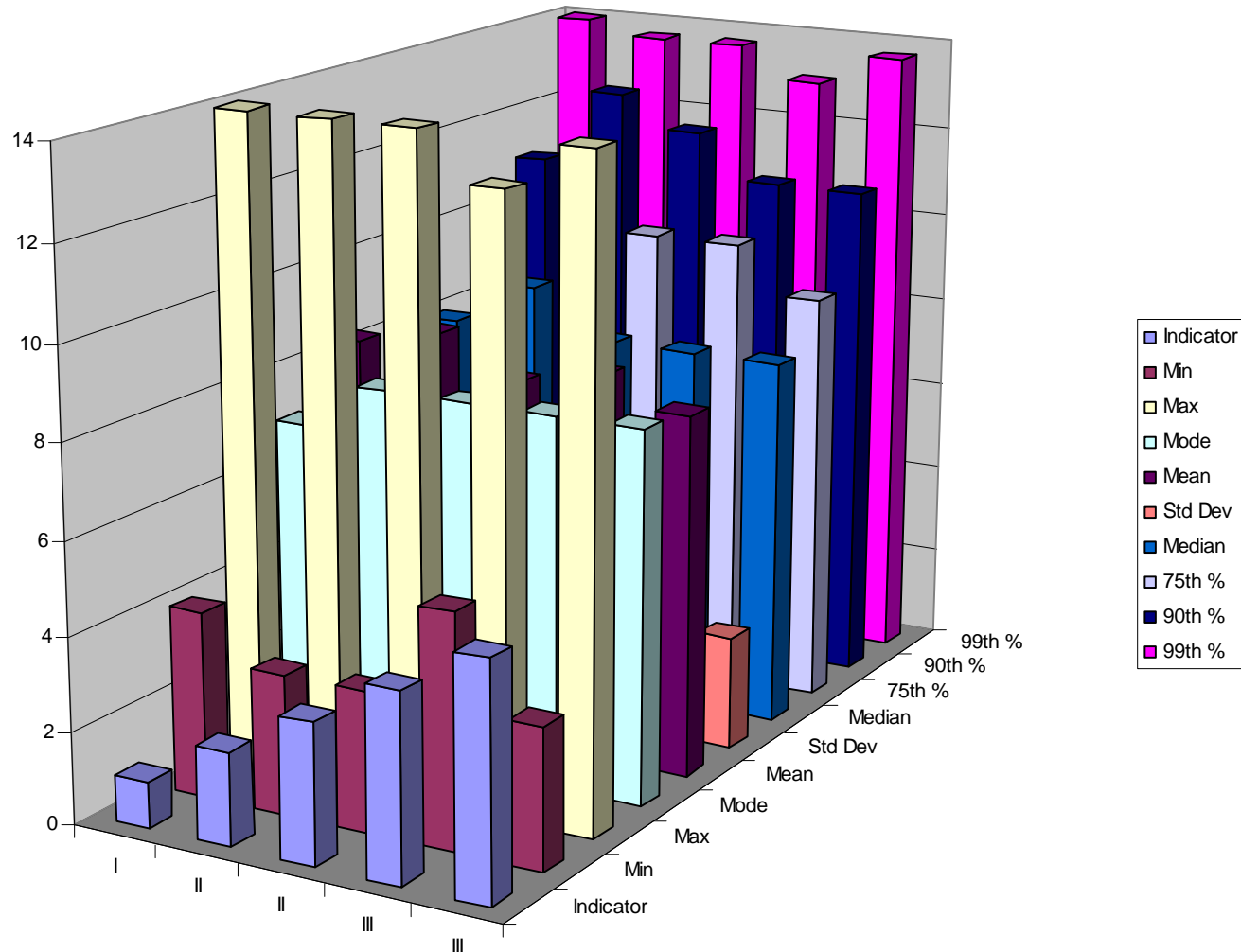
→ Combined scoring system

↳ Synthetic scores by aggregation

↳ 3 (all experts agreed that software not innovative)

↳ 15 (all experts agreed that software very innovative)

Statistics of Innovation Indicators



Dimension I

(product new to firm)

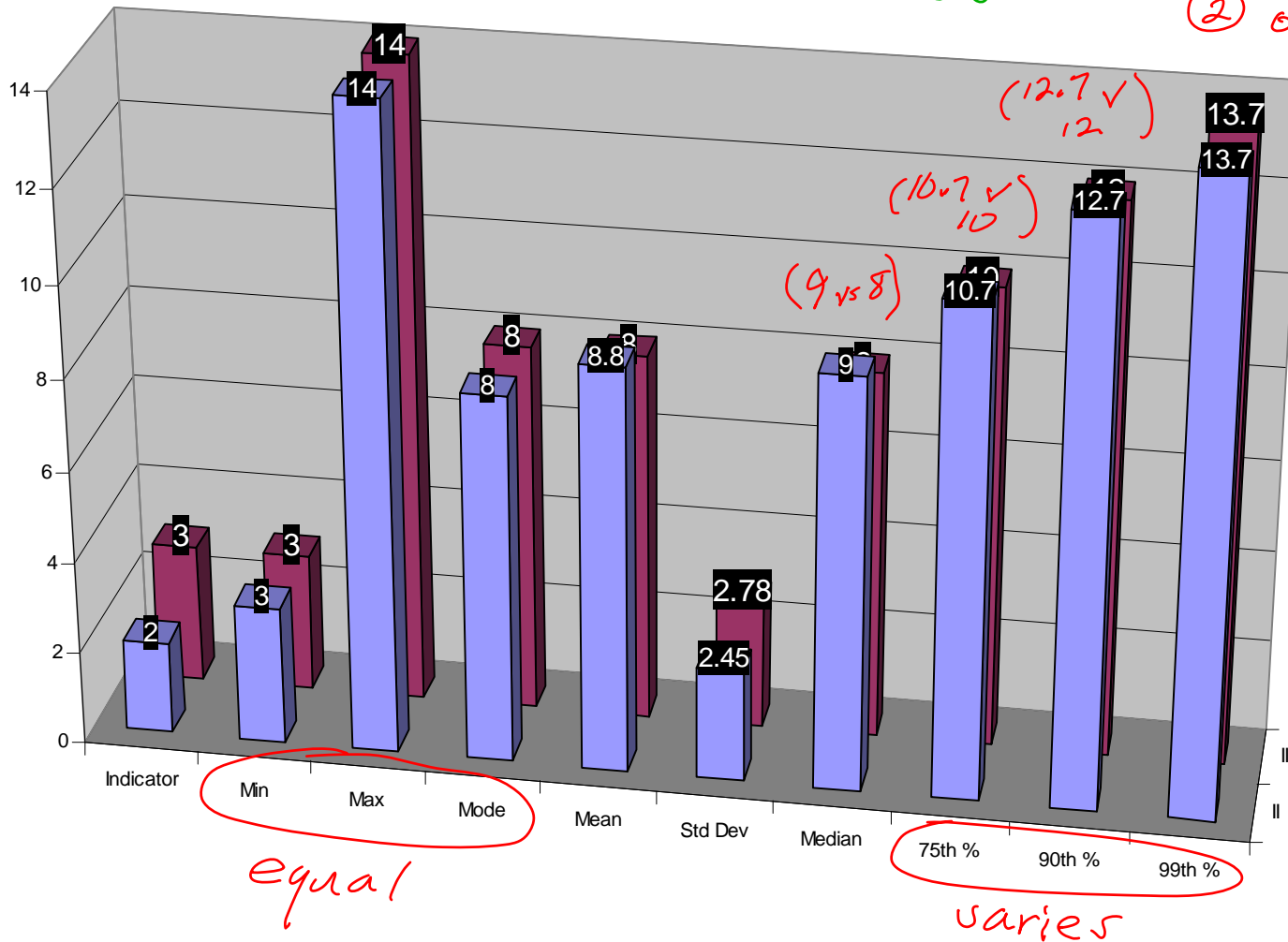


More than 50% between 6 - 8

Dimension II

- ② Satisfy User Needs
- ③ New technology

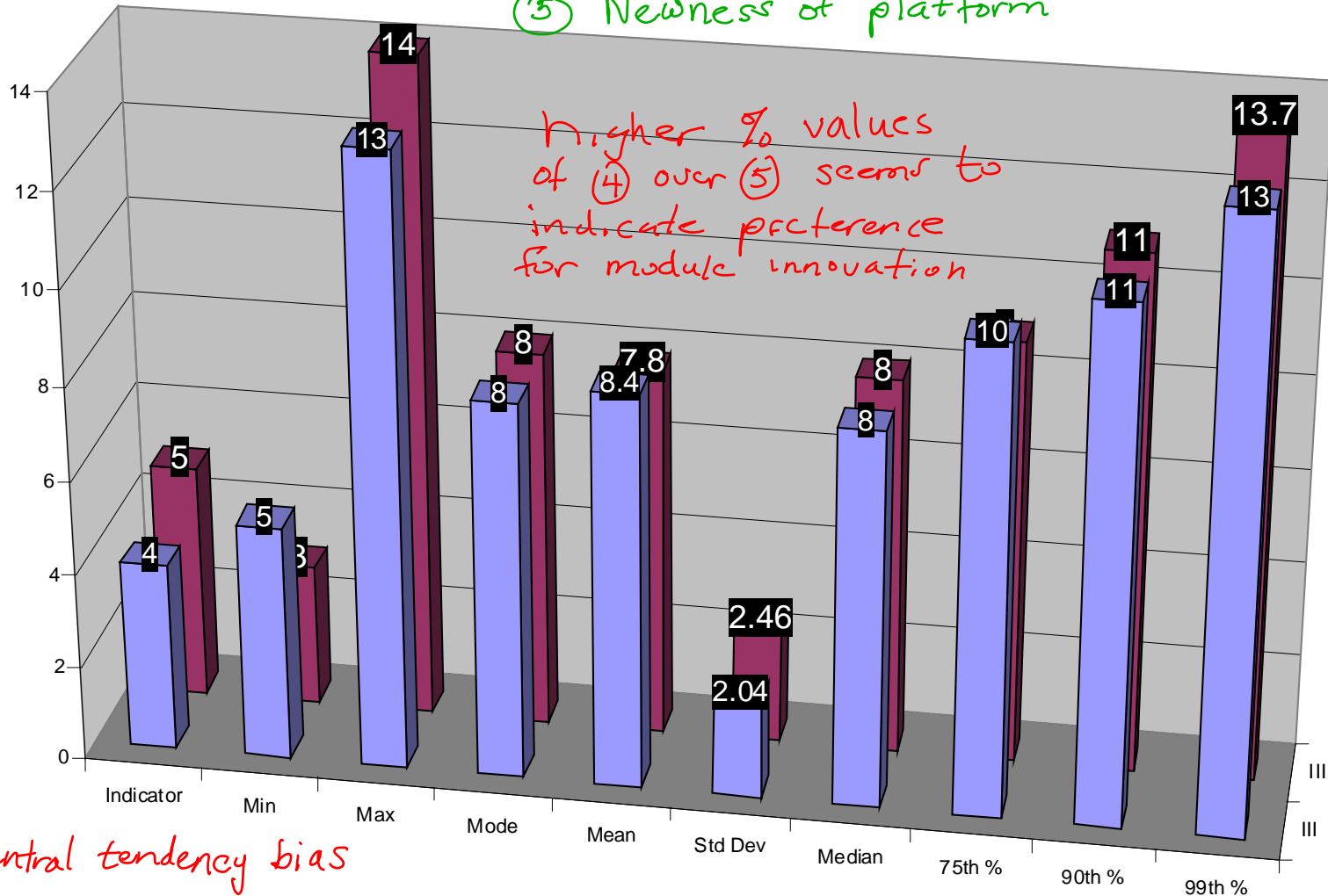
Data seems to support
② over ③



Dimension III (World)

④ Innovative modules

⑤ Newness of platform



Higher % values of ④ over ⑤ seems to indicate preference for module innovation

Central tendency bias

Indicator 5 values (5-7)

Indicator 4 values (7-9)

Summary of Data

→ Data allows insight into research questions

↳ Are software produced by Italian firms innovative?

➤ Indication that what is done is more innovative than how

↳ What are typologies of implemented innovation?

➤ Mainly incremental rather than radical

↳ Is software based on FOSS more innovative than Proprietary?

➤ 134 solutions sampled

✓ 109 proprietary

✓ 27 FOSS

➤ Overall the small subgroup of FOSS in sample received higher ratings in all 3 dimensions surveyed

Critique

→ Cons

- ↪ Authors acknowledge weakness in statistics
- ↪ Problems with size of samples
- ↪ Too small pool of “experts” surveying
- ↪ Method of selecting “experts”
- ↪ Reduced applicability of case study
- ↪ The data ultimately does not definitely answer primary question of Innovation of FOSS vs Proprietary

→ Pros

- ↪ Attempt to address important research questions in a methodical way
- ↪ The approach is interesting although not sure how so much was made out of the data

Collaborative Software Design & Development

Hipikat: A Project Memory for Software Development

Enos Jones and Danhua Shao

Overview

- Background
- Hipikat architecture
- Case studies:
 - ↳ Hipikat recommendation quality study
 - ↳ The Eclipse newcomer study
- Discussion
- Related works
- Summary

Problem

→ Challenges to newcomers:

- ↳ little knowledge about projects
- ↳ tons of artifacts: source codes, reports, web pages, etc.
- ↳ Which one is the start point?

→ Traditional solution

- ↳ mentors

→ Not effective with “virtual team”

- ↳ Time and distance separation
- ↳ Only artifacts available

→ Hipikat: the helper cat

The word belongs to the Wolof language, spoken in parts of West Africa. Hipikat means a sage or intelligent fellow, one who is in the know, and the same all-round alertness is supposed to characterise his descendants, the hep-cat, the hip cat, the hipster. (found by Enos)



Solution: Hipikat

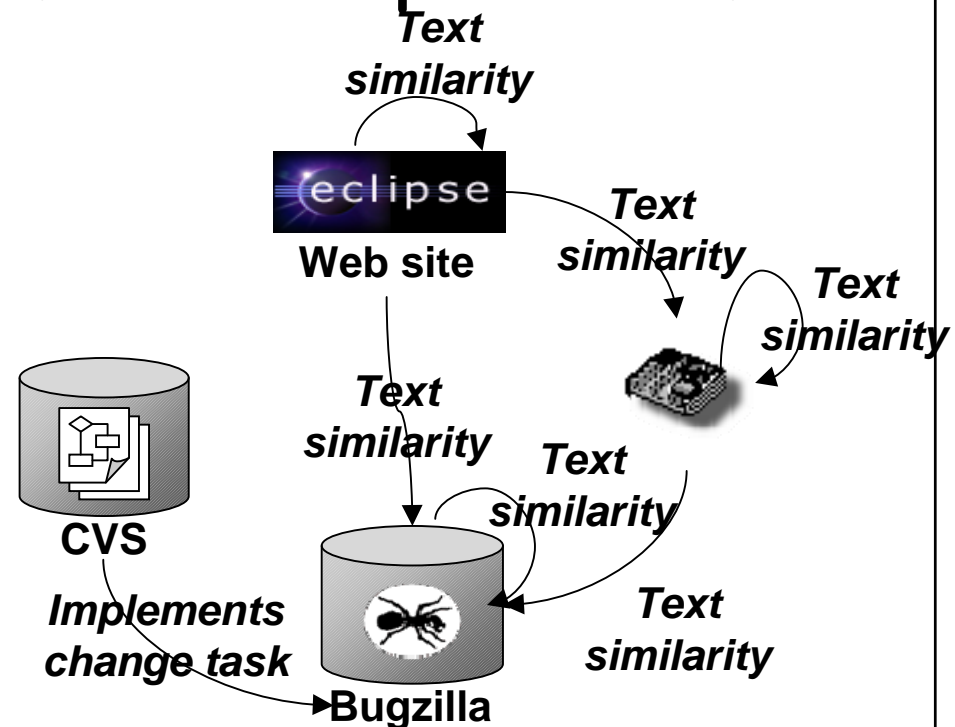
→ Form the project memory from development artifacts

↳ import artifacts:

- source code versions
- tracking system
- mailing list
- web pages
- etc.

↳ build linkages:

- Text similarity
- Implements
- etc.



→ Recommend artifacts for a query

Hipikat in Eclipse

Bug 7017 - New project location dialog misleading

Bug#: 7017 **Product:** Platform
Version: 2.0 **Platform:** PC
OS/Version: Windows 2000 **Status:** RESOLVED
Severity: normal **Priority:** P2
Resolution: FIXED **Assigned To:** ryan_cooper
Reported By: Jeff_McAffer@oti.com **Component:** UI
Target Milestone: --- **URL:**

Summary:
New project location dialog misleading

Keywords:

Description:
There have been many reports of people creating new projects and specifying default content locations getting confused and specifying the parent location rather than the location itself. For example, a typical user would create a project Foo and have it actually live in c:\projects\foo\foo

Hipikat (Bug 7017)

Type	Name	Reason	Confide...
cvs	dev.eclipse.org;/home/eclipse/org.eclipse.ui/...	Check-in close to bug resolution	High - c...
cvs	dev.eclipse.org;/home/eclipse/org.eclipse.ui/...	Check-in close to bug resolution	High - c...
bugzilla	Bug 5632 - Ability to create folders when sp...	Text similarity	0.3448...
bugzilla	Bug 7857 - Confirm project delete dialog coul...	Text similarity	0.3387

Query Hipikat

```

StorageDocumentProvider.java
* @param encoding the char
* @exception CoreException
* @since 2.0
protected void setDocument
    Reader in= null;
    try {
        if (encoding == null
            encoding= getDe
        in= new InputStream
        StringBuffer buffer=
        char[] readBuffer=
        int n= in.read(read
        while (n > 0) {
            buffer.append(re
            n= in.read(read
        }
        document.set(buff
    
```

CVS Resource History - RegistryCacheReader.java

Revision	Tags	Date	Author	Comment
1.23		5/17/02 11:2...	dwilson	16211 - Need to
1.22	v20020517 v	4/22/02 1:05 ...	dwilson	14344 - Registry
1.21		02 10:1...	johna	Remove VA/J sup
1.20		02 8:47 ...	dwilson	12364 - Registry
1.19		02 12:0...	dj	Removing .vcm_r
1.18		02 12:2...	dwilson	Remove unused l
1.17.2.1		02 12:07 ...	vlad	*** empty log me
1.17		02 5:50 AM	vlad	Removed R1.0 La
1.16		2/6/02 11:47 ...	dwilson	8288 - Registry C
1.15	v20020205	1/31/02 1:08 ...	vlad	changes for 2.0 t

Root_vk_new_build_layout 14344 - Registry
 v20020423 14346 - Registry
 v20020425
 v20020429 vk post merge

Bug 21349 X

Bug 21349 - Team error when Synchronizing with repos

Bug#: 21349 **Product:**
Version: 2.0 **Platform:**
OS/Version: Windows 2000 **Status:**
Severity: major **Priority:**
Resolution: FIXED **Assigned To:**
Reported By: shaldon@counterp.com **Component:**
Target Milestone: --- **URL:**

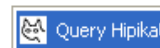
Summary:

Team error when Synchronizing with repository

Keywords:

Description:

SETUP PRODUCING THIS ERROR:
 CVSNT 1.11.1.3 is installed on "Machine A" - Server runn:



Search

File Search Help Search Hipikat Search

Search Terms

Reply from Hipikat

→ Lists relevant artifact as recommendation

Hipikat (Bug 20982)	
Type	Name
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclip
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclip
bugzilla	Bug 18667 - Allow Commit to perform Add automatically
bugzilla	Bug 22782 - [DCR] Need API to draw disbaled text in native platfor
bugzilla	Bug 10541 - We should add "Add" to the sync view
bugzilla	Bug 22372 - [CVS UI] Committing project from from Navigator witho

→ Provides reason and confidence for recommendation

	Reason	Confidence
g/eclipse/...	Bug ID in revision log	High
g/eclipse/...	Bug ID in revision log	High
	Text similarity	0.3875464
platform way	Text similarity	0.3751764

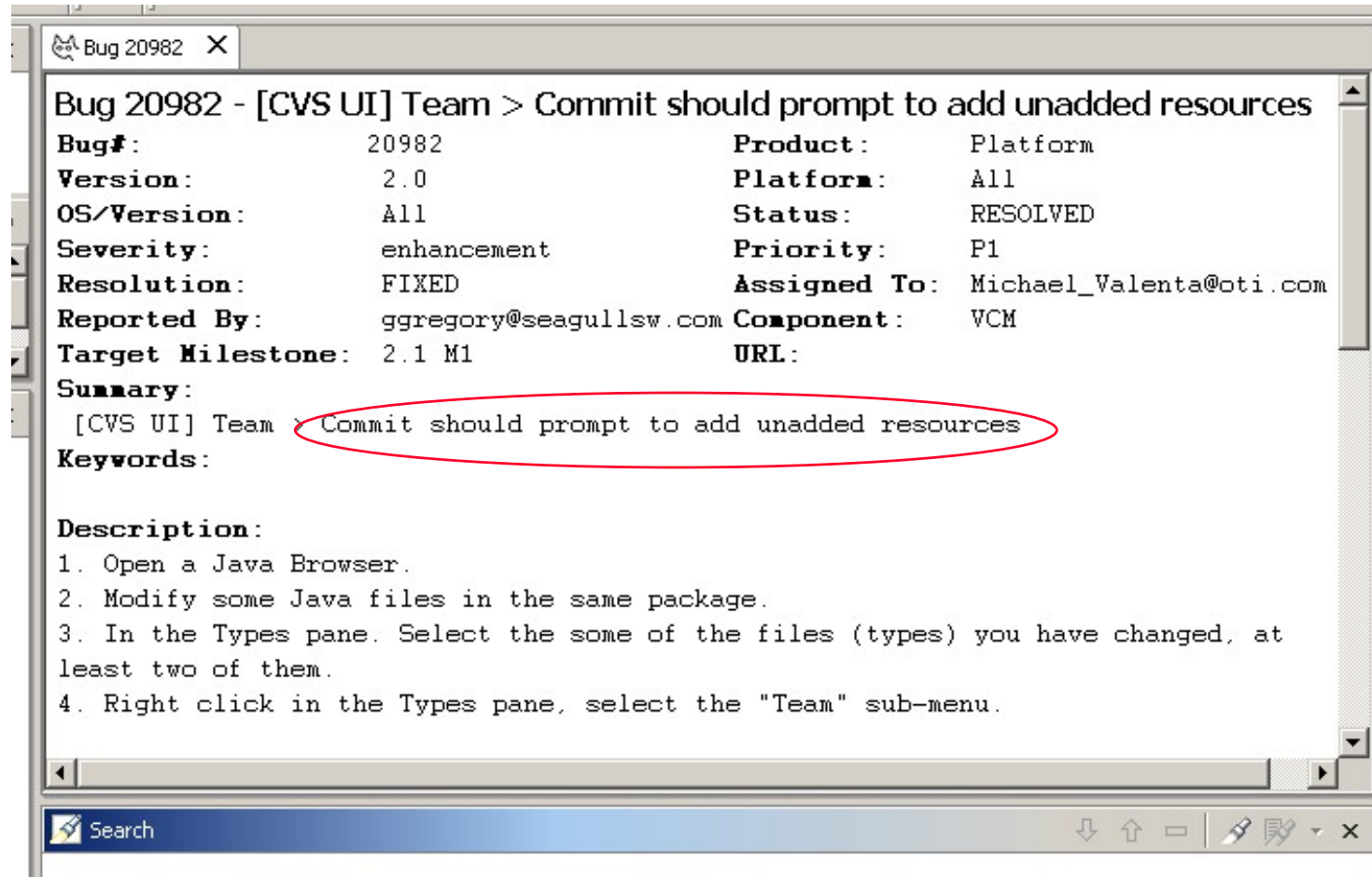
→ Can make further queries on the recommendations

	Reason
home/eclipse/org.eclipse.ui/...	Check-in clos
home/eclipse/org.eclipse.ui/...	Check-in clos
y to create folders when sp...	Text similarit
m project delete dialog coul...	Text similarit
ards] New Project: Rememb...	Text similarit
ility issues with Project Mov...	Text similarit
owse support when specifyin...	Text similarit
roject location behavior stra...	Text similarit
mal behavior in the new fold...	Text similarit

- View selected
- View Revision's Changes
- Show in Resource History
- Remove
- Remove all
- Query Hipikat

A sample scenario - 1

→ Task: fix a new bug, Bug 20982



A sample scenario - 2

- Query on artifacts relevant to new bug
 ↪ find a similar old bug, Bug 10541

The screenshot shows the Eclipse IDE interface. The top window, titled 'Bug 20982', displays the following text:

```

----- Additional Comment #7 From Kevin McGuire 2002-03-19 17:34 -----
Formerly known as: "Synchronizing new file in new package results in error"

----- Additional Comment #8 From Jean-Michel Lemieux 2002-05-09 00:27 ---
Fixed.

With some changes to the changes proposed by Kevin:

1. add action added to sync view
2. is enabled if the node or any child is a candidate to being added to cvs
3. commit doesn't touch un-added files, you can safely run commit on a parent
   folder that has added changed, exactly as in the navigator.
4. after the add action, the nodes are kept in the sync view and their
   decoration is updated to reflect their new state (this includes any parents)
5. commit no longer performs a free add, you must explicitly add/ignore
   resources first.
  
```

The bottom window, titled 'Hipikat (Bug 20982)', shows a table of search results:

Type	Name	Reason	Confidence
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/sr...	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/sr...	Bug ID in revision log	High
bugzilla	Bug 18667 - Allow Commit to perform Add automatically	Text similarity	0.387546...
bugzilla	Bug 22782 - [DCR] Need API to draw disbaled text in nati...	Text similarity	0.375176...
bugzilla	Bug 10541 - We should add "Add" to the sync view	Text similarity	0.294016...
bugzilla	Bug 22372 - [CVS UI] Committing project from from Naviga...	Text similarity	0.264780...
bugzilla	Bug 10647 - [CVS UI] Committing an untracked file does...	Text similarity	0.217704...

The 'Bug 10541 - We should add "Add" to the sync view' entry is highlighted in blue. The bottom status bar shows 'Tasks Hipikat Results Console Search'.

A sample scenario - 3

→ Query artifacts relevant to old bug

↳ find a source file: `CommitSyncAction.java`

The screenshot shows an IDE window with two tabs: "Bug 20982" and "Bug 10541". The active window displays the text of bug 10541, which includes several comments and a list of proposed changes. Below the text, a search results table is visible, showing the results of a query for the file `CommitSyncAction.java`.

Additional Comment #7 From Kevin McGuire 2002-03-19 17:34
 Formely known as: "Synchronizing new file in new package results in error"

Additional Comment #8 From Jean-Michel Lemieux 2002-05-09 00:27
 Fixed.

With some changes to the changes proposed by Kevin:

1. add action added to sync view
2. is enabled if the node or any child is a candidate to being added to cvs
3. commit doesn't touch un-added files, you can safely run commit on a parent folder that has added changed, exactly as in the navigator.
4. after the add action, the nodes are kept in the sync view and their decoration is updated to reflect their new state (this includes any parents)
5. commit no longer performs a free add, you must explicitly add/ignore resources first.

Type	Name	Reason	Confidence
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team/internal/ccvs/ui/sync/CommitSyncAction.java:1.13	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High

Tasks | Hipikat Results | Console | Search

A sample scenario - 4

→ Query the revision history of `CommitSyncAction.java`

The screenshot shows a context menu for a file named "CommitSyncAction.java" in an IDE. The menu items are:

- View selected
- View Revision's Changes
- Show in Resource History
- Remove
- Remove all
- Query Hipikat

Below the menu, a table displays revision history entries:

Type	Name	Reason	Confidence
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Bug ID in revision log	High

The IDE interface also shows a "Hipikat (Bug 10541)" tab and a "Tasks" panel with "Hipikat Results", "Console", and "Search" buttons.

A sample scenario - 5

→ Find changes on CommitSyncAction.java in fixing old bug

↪ This file can be a start point to fix the new bug

The screenshot displays the Java Source Compare tool with two versions of the `CommitSyncAction.java` file. The left pane shows version 1.13, and the right pane shows version 1.12. The code is compared line-by-line, with differences highlighted in various colors. A table at the bottom of the window provides details about the commit history.

Type	Name	Reason	Confidence
cvs	dev.eclipse.org:/home/eclipse/org.eclipse.team.cvs.ui/src/org/eclipse/team...	Check-in close to bug re...	High - checked-in within 2 minute(s)

Hipikat architecture

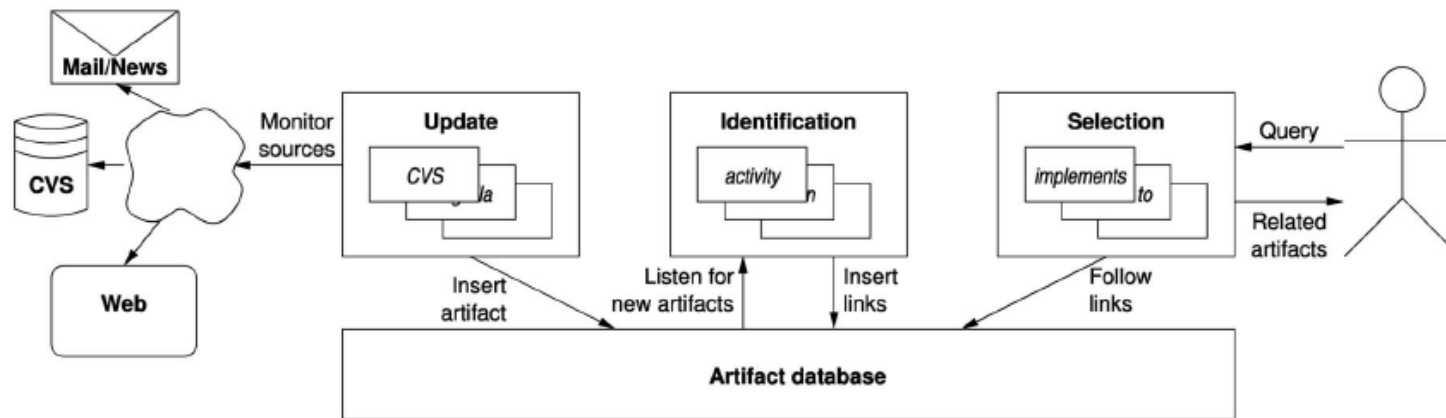
→ Client/Server structure

↳ Client:

- Query interface (Eclipse plug-in)

↳ Server:

- store and update artifacts
- identification linkage
 - ✓ log, activity, text similarity, and cvs check-in
- select candidate recommendations



Case Study 1: recommendation quality

- Goal: evaluate the effectiveness and accuracy of Hipikat recommendations
- Statistic Population
 - ↳ 215 fixed tasks in Eclipse 2.1
- Sample
 - ↳ 20
- Experiment
 - ↳ For each task
 - query Hipitak for recommendations
 - check overlap with the actual solutions in history
- Evaluation Criteria
 - ↳ $\text{precisions} = \frac{\text{correctly recommended solution}}{\text{total recommendations}}$
 - ↳ $\text{recall} = \frac{\text{correctly recommended solution}}{\text{total right solutions}}$
 - ↳ rank of the first-used recommendation

Recommendation Quality Evaluation Result

Bug id	Files		First-useful recommendation			
	Precision	Recall	Files		Constructs	
			All bugs	Completed	All bugs	Completed
23719 ^a	0.56 (5/9)	0.71 (5/7)	1	1	1	1
28382	0.50 (1/2)	1.00 (1/1)	5	1	None	
26338	0.17 (1/6)	1.00 (1/1)	2	1	2	1
30943	0.15 (4/26)	1.00 (4/4)	4	3	4	3
23321	0.13 (2/16)	1.00 (2/2)	3	1	3	1
27822	0.11 (1/9)	1.00 (1/1)	12	2	None	
19857	0.09 (2/22)	1.00 (2/2)	3	2	2	2
3248	0.07 (1/14)	1.00 (1/1)	10	4	None	
2338	0.06 (1/17)	1.00 (1/1)	7	3	10	4
9615	0.06 (1/17)	0.50 (1/2)	2	2	2	2
34641	0.05 (1/19)	0.50 (1/2)	10	4	10	4
24168	0.04 (1/24)	1.00 (1/1)	2	2	None	
20017	0.04 (1/24)	0.25 (1/4)	8	2	8	2
23365	0.04 (1/25)	1.00 (1/1)	1	1	1	1
22260	0.04 (1/26)	1.00 (1/1)	4	1	None	
28513	0.04 (1/27)	1.00 (1/1)	4	3	2	2
6732 ^b	0 (0/26)	0 (0/1)	N/A	N/A	5	1
31972	0 (0/36)	0 (0/1)	N/A	N/A	None	
32067	0 (0/14)	0 (0/1)	N/A	N/A	2	2
33182 ^c	0 (0/60)	0 (0/2)	N/A	N/A	None	
Average	0.11; 0.06	0.65; 1.00				

Case study 2: Eclipse newcomer study

→ Goal: check the effectiveness for newcomer

- ↪ Can newcomer get benefit from Hipikat?
- ↪ When and from which artifacts will newcomer query Hipikat?
- ↪ How the newcomers evaluate and utilize the recommendations from Hipikat?

→ Experiment

- ↪ two groups of participants
 - newcomers with Hipikat
 - experts without Hipikat
- ↪ two tasks
 - easy task
 - difficult task

→ Procedure

- ↪ training newcomers with Eclipse and Hipikat
- ↪ 2 interviews during development
 - middle: change plan
 - end
- ↪ Check correctness

Correctness Criteria

Correctness Criteria for the Easy Task

Hover: Displaying new properties in the hover.

- Identify `JavaLineBreakpoint` and `JavaBreakpoint` classes
- Identify markers and the message attribute
- Setting the marker message

Updating the hover: Changing the breakpoint's properties should be reflected in the hover.

- *Hover updated when user toggles suspend policy*
- *Hover updated when user changes the hit count*
- *Hover updated when user changes the condition*

Subclasses. New hover functionality should work for subclasses of `JavaLineBreakpoint`.

- *Method breakpoint hover is correct*
- *Watchpoint hover is correct*

Style: Modifications should respect existing architecture and coding style.

- Suspend policy is added in `JavaBreakpoint`, line number in `JavaLineBreakpoint`
- Strings are externalized

Correctness Criteria for the Difficult Task

Menu: Operations in the versioning context menu should be grayed-out as appropriate.

- Identify `CommitAction.isEnabled` method
- Enable the "commit" option even if some of the selected resources are new
- *Do not enable the "commit" option if all resources are in `.cvsignore`*

Detecting selected new files: When the "commit" option is selected in the versioning context menu, the code should check for the existence of any new files in the active selection.

- Identify `CommitAction.execute` method
- Look for new resources in the active selection
- *Check in selected subdirectories for new resources (any number of levels deep)*

Message dialog: If there were any new resources in the active selection, a dialog should be put up to ask the user whether to add them to CVS management before proceeding with the commit.

- Put up the dialog if any of the selected resources are new
- If the user presses OK, add the new resources to CVS management
- *If the user presses No, do not do anything to the new resources, but proceed with the commit process*
- *Cancel aborts the whole commit process*

Committing new resources: If the user selected "yes" or "no" in response to the message dialog, the rest of the commit process should proceed appropriately.

- Show the check-in comments dialog
- All selected managed resources are committed to the repository (including the new ones)

Style: Modifications should respect existing architecture and coding style.

- Do not simply call `ForceCommitSyncAction` to display the dialog (increases coupling)
- Externalize message in the dialog
- Use the *visitor* pattern to look for new resources in the selected subdirectories

Newcomer evaluation results

→ The easy task

- ↳ Experts performed worst on handling the special cases
- ↳ 25% of experts got right solution
- ↳ 38% of newcomers got right solution
- ↳ 63% of newcomers got basically right solution

→ The difficult task

- ↳ 75% of experts solved it correctly
- ↳ 38% of newcomers solved it correctly

→ Access to Hipikat

- ↳ more query were made on the difficult task
- ↳ almost all the queries were made in first hour

→ Evaluation to Hipikat recommendations

- ↳ are used for code re-use or change positions identification
- ↳ in difficult task, are not used for code understanding

Discussion

→ Model

- ↳ Recommendation granularity
 - current: file level
- ↳ Time awareness
 - some data may get out-of-date
- ↳ Collaborative recommendations
 - need rating from users

→ Implementation

- ↳ Presentation of query results
 - flat list or cluster
- ↳ Scale up

→ Applicability

- ↳ Refine existing functionality
- ↳ Learn API usage
- ↳ Recovering design rational
- ↳ Avoid common errors

Related works

→ Group/organize memory

↳ Design Assistant

- recommend design decisions, code structures, and usage of API
- need experts to construct project memory manually

↳ TeamInfo

- artifact: email
- linkage: keywords

→ Recommender systems

↳ Remember Agent & CodeBroker

- single artifact source: email or source code

→ Programming by examples

↳ Recommend source code examples from existing projects or intentionally constructed library

→ Mining software repository (versions)

↳ CVSSearch

↳ Expertise Browser

↳ Expertise Recommender

Summary

- In most cases, Hipikat can provide useful recommendations.
- With the help from Hipikat, newcomers can achieve comparable results to experts.
- Questions:
 - ↳ the recommendation quality depends on the amount of artifacts
 - ↳ the effect of developers' practice on the quality of recommendations
 - ↳ shallow understanding of recommendations

Pros/Cons

→ Pros

- ↳ well organized study
- ↳ experiment with real projects
- ↳ show both benefit and shortcomings

→ Cons

- ↳ **construct validity**
 - **criteria in solution evaluation**
 - ✓ Only count the solution in real project
 - ✓ might be multiple solutions for the same task
- ↳ **internal validity**
 - **Eclipse newcomer study**
 - ✓ done: newcomer with Hipikat vs. expert without Hipikat
 - ✓ suggest: newcomer with Hipikat vs. newcomer without Hipikat

Reference

- Hipikat: A Project Memory for Software Development (TSE June 2005, Vol. 31, No. 6)
- Hipikat: recommending pertinent software development artifacts (ICSE 2003)
- Hipikat project:
<http://www.cs.ubc.ca/labs/spl/projects/hipikat/>

Questions/Comments

→ ?/!