Integration of Open Source Software Projects in IT Education

Funded by

National ICT R&D Fund
Ministry of Information Technology
Software Engineering Education
Boehm 1998

• The Stake Holders
  – Teaching of software engineering requires satisfying the needs of many stakeholders including the students, the industry recipients of the students, software engineering community as parties interested in improved practices, and the instructors and teachers.

• Desirable Knowledge and Skills – Student/Employers perspective
  – long half-life fundamental knowledge,
  – relevant and marketable near-term experience and skills, and
  – full-spectrum coverage of software development life cycle and processes
Challenges

- Software engineering techniques apply to solving large-scale problems involving several team members and roles.
- Limitation of the classroom setup
  - projects that really require application of software engineering principles can neither be discussed in the class nor given as assignments.
  - When applied to small problems, SE techniques appear exaggerated and students find it difficult to appreciate the significance of software engineering principles.
Final Year Capstone Projects

• Typically two semester long capstone project.
• Attempts to create a realistic experience of professional software development in classroom setting.
• Weakness:
  – Emulation of customer role, and
  – Lack of the rigor required for developing large industrial quality software products.
• Results in only a fractional fulfillment of the objective.
• Requires longer than desired training period for new graduates once they are hired by the industry.
Solution

- Engage the students in real projects with real clients.

- Problems:
  - Availability of real projects and real clients
  - Sustainability of such efforts
  - Synchronization of timelines of industry projects with semester schedule

- Proposed Solution:
  - Participation of students in developing open source software
Open Source Development Process

Figure 1: Open Source Development Cycle
Open Source Development Process

- Quality of the open source software depends largely upon the quality of the review process.
- From a software engineering perspective, an open source project involves requirement analysis and design, coding, unit testing, peer review, and configuration management.
- Can Open source provide enough projects to sustain?
## Open Source – Some Statistics

<table>
<thead>
<tr>
<th>Topic</th>
<th>Number</th>
<th>Topic</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering</td>
<td>528</td>
<td>Religion and Philosophy</td>
<td>346</td>
</tr>
<tr>
<td>Database</td>
<td>7241</td>
<td>Scientific Engineering</td>
<td>15725</td>
</tr>
<tr>
<td>Desktop Environment</td>
<td>3910</td>
<td>Security</td>
<td>3400</td>
</tr>
<tr>
<td>Education</td>
<td>5211</td>
<td>Sociology</td>
<td>444</td>
</tr>
<tr>
<td>Formats and protocols</td>
<td>3109</td>
<td>Software Development</td>
<td>29643</td>
</tr>
<tr>
<td>Games &amp; Entertainment</td>
<td>18035</td>
<td>Systems</td>
<td>23920</td>
</tr>
<tr>
<td>Internet</td>
<td>29569</td>
<td>Terminals</td>
<td>711</td>
</tr>
<tr>
<td>Multimedia</td>
<td>16172</td>
<td>Text Editors</td>
<td>3325</td>
</tr>
<tr>
<td>Office/Business</td>
<td>10200</td>
<td>Other</td>
<td>2678</td>
</tr>
<tr>
<td>Printing</td>
<td>530</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Number of listed projects under each topic at Sourceforge.net as on May 10, 2007

Lots of domains – lots of project in each domain
Open Source – Some Statistics

Distribution of top 500 projects listed on Sourceforge.net as of 28/8/07

- 130 projects in teams of 1-2
- 96 projects in teams of 3-5
- 100 projects in teams of 6-10
- 55 projects in teams of 11-15
- 104 projects in teams of 16-25
- 47 projects in teams of 26-50
- 18 projects in teams of 51-100
- 5 projects in teams of 100+

Lots of active projects with decent team size
Proposed Model

- In the sixth semester, selected students will be asked to participate in the project.
  - "Open" projects available in the OS repository will be chosen based on a combination of factors such as team-size, maturity, expertise of SRO's, current and future application trends.
- First Six Months
  - understanding the problem domain under the supervision of their SROs and ROs
  - identify and fix any defects that may be found in the existing code base.
- Next Six Months
  - contributing to the project code, adding new functionality and addressing any open issues with the product.
- Next Six Months
  - The best of the students would then be identified and picked as mentors for the next cohort entering their project year.
  - The self-sustaining cycle would then continue.

<table>
<thead>
<tr>
<th>6th Semester</th>
<th>7th Semester</th>
<th>8th Semester</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain understanding and bug fixing</td>
<td>Development</td>
<td>Mentoring</td>
</tr>
<tr>
<td>Top students only</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Top students only
Org Chart for the Project

Dr. Naveed Malik
Co-PI
VU

Dr. Fakhar Lodhi
PI
NU

SE Expert
(for curriculum development)

Senior RO

Dr. Mansoor Sarwar
Co-PI
PUCIT

Senior RO

Senior RO

Senior RO

RO

RO

RO

RO

RO

RO

RO

RO

RO

Student Participants

Student Participants

Student Participants
Time (months) – total duration: two and half years

Initial study and Model development
- Initial study
- Participation in the projects and model development
- Fine tuning the model

Execution, fine tuning and release
- Execution, Analysis, synthesis, and training
- Mentoring and monitoring

Core team

Students
- Debugging
- Development
- Mentoring

Activities of the first batch of students in their 6th, 7th, and 8th semesters respectively (summer included)

Activities of the second batch of students in their 6th and 7th semesters respectively (summer included)

Development and execution of the model
Benefits

- Understanding of the problem domain
- Exposure to professionally written code.
- Feedback from professionals
- Interaction with global teams – improved communication and interaction skills
- Development practices in a distributed environment
- Ethical practices and respect of intellectual property
- Confidence building
- Positive contribution in development of open source software
- May encourage students with entrepreneurial mindset to take the initiative
- Higher market demand
- Less training requirements
- May result in initiating indigenous OSS projects
Challenges

- Lots of unknowns
- Finding mentors
- Identifying suitable projects
- Motivating students
Batch 1

- 20 Students
- OpenBravo – ERP
- OrangeHRM – HR Management System
- Activities
  - Domain Understanding
  - Bug Fixing
  - Feature Development
Openbravo

- Bug fixes contributed: 4
- Bug fixes worked on: 9
- Feature requests submitted: 2
- Feature requests worked on: 5
- Customizations:
  - SMS Alerts (Iteration 1 completed)
  - Mobile Sales ‘n’ Stocks Application (Iteration 1 completed)
Orange HRM

- Bugs / Feature Request contributed: 8
- Bugs / Feature Requests active: 5
- Bugs / Feature Requests worked on: 13
Lesson Learnt

- Domain understanding
- Maintaining the rhythm
- Community interaction
- Customization
Acknowledgement

• GUNI – Global University Network for Innovation

The GUNI was set up by UNESCO, the United Nations University (UNU) and the Technical University of Catalonia (UPC), which hosts the Secretariat. One of GUNI's activities is the Observatory of Good University Practices, which aims to identify innovative experiences related to the social commitment of universities. The Observatory is a point of reference and a place for dialogue and exchange for those interested in the social commitment of universities.

We would like you to supply us with information on your experience through the OpenBravo program. From our point of view this good practice is an important referent in promoting innovation in the field of higher education, and we thought that it can be a reference to universities around the world.

For this reason we ask for your collaboration. We shall be grateful if you could provide us with more information regarding the experience that your institution has carried out, the needs covered by OpenBravo and the benefits that have been produced to your University.
Batch 2

- 18 students
- Evenly divided in two groups
- Started in Feb. 2009
- First 3-4 months were spent in basic training and domain understanding
- Two of the students are participating in GSoC
What is GSoC?
Google is working with several open source, free software, and technology-related groups to identify and fund several projects over a three month period. Historically, the program has brought together nearly 2,500 students with over 180 open source projects to create millions of lines of code.

What is the purpose?
- Get more open source code created and released for the benefit of all.
- Inspire young developers to begin participating in open source development.
- Help open source projects identify and bring in new developers and committers.
- Provide students in Computer Science and related fields the opportunity to do work related to their academic pursuits (think "flip bits, not burgers").
- Give students more exposure to real-world software development scenarios.
Project Introduction
Change Tracking of Data and Structure

- This Project is related to **MySQL and its official web-client phpMyAdmin**
  - Idea is to add “Change Tracking/Synchronization module”.
  - The functionality would help the developer of phpMyAdmin to migrate changes from one database to another by just pressing the synchronization button.
  - Possible changes could be related to
    - adding/removing/updating tuples,
    - creating/dropping tables,
    - changing schema and maintaining table evolution from state A to B.
Key Information

- Mentor
  - Michal Cihar (http://blog.cihar.com/)
- Project Page
  - http://pmapplanet.cihar.com/
- Duration
  - 3 months
- Stipend
  - $4500
Implementations in Industry
OrangeHRM

- @CambridgeDocs (pvt) Ltd
  - Proposal for implementation (Done)
    - Time & Attendance
    - Personal Information Management
  - Bug Fixing & Feature Requests (in process)
  - Next Targets
    - Attendance through Biometric Device
    - Recruitment Module
- @Ascertia (pvt) Ltd
  - Proposal for implementation (in process)
- Pronto
  - Proposal for implementation (in process)
Openbravo Implementation

- Discussions going on with Pronto Promotions
  (http://www.prontopromo.com)
- Building scenarios to evaluate the process and the way Openbravo would support the process
Other Activities

- Make OrangeHRM accessible over Web Services
- Performance Review Module
- Implementation & Client Requests
  - Attendance & TimeSheet Reports
  - Attendance through Biometric Device
- Make user rights configurable
- iFone: Time, Attendance & leave
- Implementations & Client Requests
Other Activities

- Openbravo ERP
  - SMS Alerts
  - Mobile Sales ‘n’ Stocks Application

- Openbravo POS
  - Customer Loyalty Management
  - Employee Presence Management

- Bridging OB and OHRM
# IMPLEMENTATION, RESULTS, DISCUSSION, AND OBSERVATIONS

<table>
<thead>
<tr>
<th>Projects</th>
<th>Student Groups</th>
<th>No. Of Students</th>
<th>Average CGPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Capstone Projects</td>
<td></td>
<td>20</td>
<td>2.95</td>
</tr>
<tr>
<td>OrangeHRM</td>
<td></td>
<td>4</td>
<td>2.96</td>
</tr>
<tr>
<td>phpMyAdmin</td>
<td></td>
<td>2</td>
<td>3.47</td>
</tr>
<tr>
<td>Open Source Projects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Openbravo ERP</td>
<td></td>
<td>8</td>
<td>2.93</td>
</tr>
<tr>
<td>Openbravo POS</td>
<td></td>
<td>4</td>
<td>2.98</td>
</tr>
</tbody>
</table>
Technical Areas

- Requirement Engineering
- Design
- Coding
- Debugging
- Unit Testing
- Code review
- Exposure to new tools and technologies
- CM and version control
The Project, Team, and Personal Activities

- Project Management
- Time Logging
- Community/client interaction/ Communication Skills
- Organizational behavior
- Team Work
- Time management
Performance Comparison
Team performance in technical areas

TCP | OrangeHRM | phpMyAdmin | OpenBravo ERP | OpenBravo POS | Average
---|---|---|---|---|---
1.4 | 2.8 | 2.9 | 1.7 | 2.2 | 2.4

Team Performance in Project, Team, and Personal Activities

TCP | OrangeHRM | phpMyAdmin | OpenBravo ERP | OpenBravo POS | Average
---|---|---|---|---|---
0.7 | 3.1 | 3.3 | 2.7 | 2.9 | 3.0
Observations

• The relative learning of students in our project and those taking traditional capstone project, our students did better in both technical and non-technical areas.

• The difference is much broader in the non-technical areas as compared to the technical areas. In fact, the TCP students have negligible achievement in non-technical areas. Out of the six non-technical areas, TCP students have almost nothing to show except for team work.
Observations

- Performance of the TCP students is better than most IOPS students in design related activities.

- In areas involving requirement, coding, and new tools and technologies, the performance of both groups are comparable.

- In the rest of the areas (debugging, unit testing, code review, and CM and version control) the TCP students have almost nothing to show.
• In both the technical and non-technical categories, performance of the 4 IOSP projects is ranked in the following order: phpMyAdmin, OrangeHRM, OpenBravo POS, and OpenBravo ERP.
Other observations and Lessons Learnt

• Factors related to
  – Time commitment
  – Project complexity
  – Community feedback
  – Group formation
  – Bug fixing vs feature development
Proposed Refinements in the Model

- To lessen the barrier to entry, students should be engaged in projects which are of relatively smaller size and complexity.

- In order to reduce dependence on community feedback and maintain a proper project rhythm, it is proposed to identify vertical components in the adopted projects and initiate work on these components indigenously.

- Participation in the project during summer should be made mandatory as that is the time when students can work on the assignment with dedicated time commitment.
FUTURE WORK

• In future, we want to expand the scope of the experiment and include more students and institutes in this exercise.
• Using students as mentors has not been experimented with – we plan to do that in the next semester
• Analyze results from the VU and PUCIT.
• VU is of special interest as it uses a distance learning model with very little direct one-to-one interaction with its students.
Thank you! Questions?