Concordia University
Department of Computer Science and Software Engineering
SOEN 691D/2 DD Open Source Software and Mining (4 CREDITS)
COURSE OUTLINE
Peter C. Rigby
Fall 2013

**Course Description**

Development practices and software tools (*e.g.* git, Eclipse) used in Open Source Software (OSS) projects. Comparisons of OSS practices with Agile and traditional (*e.g.* CMMI) styles of development. Human collaborative aspects of software development. OSS business models (or how to make money off free software). OSS licensing models and their impact on use in commercial/private software. Mining software repositories (version history, email discussion, etc) to predict defects. The group project will involve the maintenance/evolution of an existing OSS project. Students will interact with professional OSS developers. A short, group paper will replicate results from the MSR community.

*Prerequisite:* SOEN 6431 or equivalent

**Grading Scheme**

Grades will be based on a software practices exam, a group project involving fixing a bug on an OSS project and describing the practices used, and a mining software repositories paper (in groups or alone) which will involve analysis using machine learning (*e.g.* J48, Naive Bayes) or the creation of a statistical model (*e.g.* a regression) of the data.

| | |
|---|---|
| Software Practices Exam | 40% |
| Bug fix on an OSS Project | 20% |
| In class discussions and summaries of papers | 15% |
| Research paper including model of data | 25% |
| Total | 100% |

You must pass the exam to pass the course. Students wanting to do a larger research project may be allowed to transfer the weight of the bug fix project to the research paper.

**Textbook and References**

The course textbook will be "Producing Open Source Software" by Fogel. It is freely available for download at `http://producingoss.com/`

Readings on Mining Software Repositories will be selected during the class based on student interest.

**Professor**
Peter C Rigby, PhD
Office: EV 3.118
Telephone: 514-848-2424 x3167
Email: peter.rigby@concordia.ca (please include SOEN 691D in the subject of all emails)

Office hours: Friday 3:30 4:30 or by appointment

**Lectures**
There will be one lecture per week on Friday in *Room H-523 from 5:45 to 8:15.* Please see the next page for detailed schedule.

There will be two guest lectures by developers working on OSS projects. One will be from a Google the other will be from a FreeRDP developer.

Each week we will first discuss OSS practices and following the break we will discuss research papers on mining software repositories.

Moodle Course website: Frequently check the website for the schedule, announcements, course material, projects, etc.

IMPORTANT NOTE: In the event of extraordinary circumstances beyond the University's control, the content and/or evaluation scheme in this course is subject to change.

**Academic Honesty**
Violation of the Academic Code of Conduct in any form will be severely dealt with. This includes copying (even with modifications) of program segments. You must demonstrate independent thought through your submitted work. The Academic Code of Conduct is available at: `http://web2.concordia.ca/Legal_Counsel/policies/english/AC/Code.html#`

| | Software Practice Topics |
|---|---|
| 1 | Introduction. What is OSS and case examples. Choosing a project. |
| 2 | "The Cathedral and the Bazar". OSS vs Agile and traditional development approaches |
| 3 | OSS practices, including releasing, testing, and review |
| 4 | Patching, branching models, and software development tools. Subversion, git, Eclipse, Bugzilla, gerrit, ReviewBoard, etc |
| 5 | Overview of software maintenance and evolution |
| 6 | The collaborative aspects of software development and "How To Ask Questions The Smart Way"[1] |
| 7 | OSS Licenses and business models. How to make money from free software and avoid lawsuits from incorrectly using OSS software. |
| 8 | Software Practices Exam |
| | Mining Software Repository Topics |
| 1 | Introduction to mining software repositories. Software measures and archival data |
| 2 | Cleaning data, understanding assumption of statistical tests, and interpreting the results: data mining using Weka and R |
| 3 | Interpretation of results: validation and triangulation |
| 4 | Writing up scientific results |
| 5 | Project Presentations. Submission of final paper and patches submitted to OSS project |

**Topics**

The above table represents possible topics to be covered in class. These will vary dependent on the interest of students.