

**Carleton University**  
**School of Computer Science**  
**COMP 5117: Mining Software Repositories**  
**Fall 2020**

## **Class Schedule**

Seminars are held every Tuesday from 11:35 AM to 2:25 PM (online).

Slack is used for course communication, news and reminders. Please join [COMP5117 channel](#).

Schedule of seminars is posted on the course website.

Fall break: October 26–30, 2020 (no classes).

## **Instructor**

Dr. Olga Baysal

Email: [olga.baysal@carleton.ca](mailto:olga.baysal@carleton.ca)

Office: HP 5414

Office Hours: by appointment only

Website: <http://olgabaysal.com/>

## **Course Website**

[http://www.olgabaysal.com/teaching/fall20/comp5117\\_f20.html](http://www.olgabaysal.com/teaching/fall20/comp5117_f20.html)

## **Short Description**

Introduction to the methods and techniques of mining software engineering data. Software repositories and their associated data. Data extraction and mining. Data analysis and interpretation (statistics, metrics, machine learning). Empirical case studies.

## **Description**

Software development projects generate impressive amounts of data. Mining software repositories research aims to extract information from the various artifacts produced during the evolution of a software system and inferring the relationships between them. This course will introduce the methods and tools of mining software repositories and artifacts used by software developers and researchers. Students will learn to extract and abstract data from software artifacts and repositories, such as source code, version control systems and revisions, issue-tracking systems, and mailing-lists and discussions. Students will also learn about various techniques of analyzing this data in order to identify meaningful relationships, patterns and trends, to recover behaviours and software development processes from evidence, or to empirically test hypotheses about software development.

## Prerequisite

Students are expected to have some background in software development and software engineering. Prior knowledge of data mining, machine learning, statistics and natural language processing would be an asset, but is not required.

## Objectives

This graduate course explores the mining of data in software repositories in order to help researchers gain empirically based understanding of software development practices, and to support practitioners in managing, maintaining, and evolving complex software projects. The course will discuss leading research in the areas of mining software repositories. Papers discussed in this course will give students a glimpse of leading research which transforms software repositories from static record keeping repositories to active repositories that are used by researchers and practitioners to better understand and predict software development activities instead of depending on personal experiences and intuition. Students will be able to extract and analyze information from multiple software repositories in order to reason about existing software systems and development processes, as well to validate hypotheses about software development using data extracted from existing software systems.

## Content Overview

The course will be adjusted according to students' interests and experience. This is an overview of the kinds of topics the course could cover:

- Mining software repositories
- Development team processes
- Software development tools and environments
- Software analytics
- Software visualization
- Sentiment analysis
- Machine learning for SE
- Mining social data
- Collaborative development
- Quantitative and qualitative evaluation of software engineering research

## Evaluation

- Weekly paper reviews: 10%
- Class participation and discussion: 20%
- Paper presentation: 10%
- Course project: 60% (10% project presentation + 50% project report)

## University Policies

For information about Carleton's academic year, including registration and withdrawal dates, see [Carleton's Academic Calendar](#).

**Pregnancy Obligation.** Please contact your instructor with any requests for academic accommodation during the first two weeks of class, or as soon as possible after the need for accommodation is known to exist. For more details, visit [Equity Services](#).

**Religious Obligation.** Please contact your instructor with any requests for academic accommodation during the first two weeks of class, or as soon as possible after the need for accommodation is known to exist. For more details, visit [Equity Services](#).

**Academic Accommodations for Students with Disabilities.** If you have a documented disability requiring academic accommodations in this course, please contact the Paul Menton Centre for Students with Disabilities (PMC) at 613-520-6608 or [pmc@carleton.ca](mailto:pmc@carleton.ca) for a formal evaluation or contact your PMC coordinator to send your instructor your Letter of Accommodation at the beginning of the term. You must also contact the PMC no later than two weeks before the first in-class scheduled test or exam requiring accommodation (if applicable). After requesting accommodation from PMC, meet with your instructor as soon as possible to ensure accommodation arrangements are made. For more details, visit the [Paul Menton Centre](#) website.

**Survivors of Sexual Violence.** As a community, Carleton University is committed to maintaining a positive learning, working and living environment where sexual violence will not be tolerated, and survivors are supported through academic accommodations as per Carleton's Sexual Violence Policy. For more information about the services available at the university and to obtain information about sexual violence and/or support, visit: [carleton.ca/sexual-violence-support](http://carleton.ca/sexual-violence-support)

**Accommodation for Student Activities.** Carleton University recognizes the substantial benefits, both to the individual student and for the university, that result from a student participating in activities beyond the classroom experience. Reasonable accommodation must be provided to students who compete or perform at the national or international level. Please contact your instructor with any requests for academic accommodation during the first two weeks of class, or as soon as possible after the need for accommodation is known to exist. For more details, see the [policy](#).

**Student Academic Integrity Policy.** Every student should be familiar with the Carleton University student academic integrity policy. A student found in violation of academic integrity standards may be awarded penalties which range from a reprimand to receiving a grade of F in the course or even being expelled from the program or University. Examples of punishable offences include: plagiarism and unauthorized co-operation or collaboration. Information on this policy may be found [here](#).

**Plagiarism.** As defined by Senate, "plagiarism is presenting, whether intentional or not, the ideas, expression of ideas or work of others as one's own". Such reported offences will be reviewed by the office of the Dean of Science.

**Unauthorized Co-operation or Collaboration.** Senate policy states that "to ensure fairness and equity in assessment of term work, students shall not co-operate or collaborate in the completion of an academic assignment, in whole or in part, when the instructor has indicated that the assignment is to be completed on an individual basis". Please refer to the course outline statement or the instructor concerning this issue.