



CS 6474 Social Computing: Introduction

Munmun De Choudhury

munmund@gatech.edu

Week 1 | August 22, 2016

Part I: Course Structure and Information

Learning Objectives

- Course Website:
http://www.munmund.net/CS6474_Fall2016.html
- **Goal: How and why social computing works?**
 - What's the right way to design social computing systems? What sense can you make out of all the data people leave behind? What's still out there to infer, understand, and build?
- At the completion of this course, you'll be able to:
 - Understand important features of social computing
 - Design and prototype new social computing systems
 - Analyze data left behind in social media
 - Understand the research issues (theoretical and practical) in this field

Grading

- Responses to Class Readings (on Piazza) - 15%
- Assignment I - 15%
- Assignment II - 15%
- Class Participation - 10%
- Term Project - 45%
 - Project Proposal - 5%
 - Midterm Project Presentation - 5%
 - Midterm Report - 10%
 - Final Project Presentation - 5%
 - Final Report - 20%

Required Skills

- *Technical*: any object-oriented/scripting language like Python, Perl, C#; some frontend development skills/web programming skills (ajax, javascript, php) if your project is about building a social tool
- *Statistics/machine learning*: preliminary knowledge of working with some data and using some analytical software
- *Reading/writing*: approximately two papers assigned for each class (four in a week); significant weightage on weekly class readings, mid-term and final term project report

Course Materials/Logistics

- Links to papers assigned to classes will be available on the course website.
 - To access ACL Digital library links to papers, use GT VPN.
 - Paper assignments are subject to change, so always check the online schedule before you read and write your reflections.
- No official text books.
- Use of open source libraries encouraged.
- Assignments will be released 3-4 weeks ahead, on T-Square.
- Lecture slides of each class will be made available on the course website within a week's time.
- Classes will involve time to time in-class activities – encouraged to bring laptops, but strictly restricted to classroom needs and use.

Participation in Class Readings

- Write short blurbs on Piazza (under “reflections”) about the readings assigned for a particular class
 - Blurbs can range from 300-600 words in length
- Blurbs should focus on the following, but not limited to:
 - What is the problem the paper is studying? Why is it important?
 - How does the paper relate to the topics being covered in the class?
 - What is the novelty of the paper?
 - Ways the work in the paper could improve
 - Why the particular method/data used in the paper was appropriate

Piazza

- Sign up with your GT information at this link [IMPORTANT]:
piazza.com/gatech/fall2016/cs6474
- Class link:
piazza.com/gatech/fall2016/cs6474/home

Assignment I

- **Due: October 19, 2016**
- Questions will focus on the topics of the reading materials covered until the point the assignment is released.
- Design focused questions – **Option A:** develop a mockup
- Design focused questions – **Option B:** build a web application, browser extension etc.

- What to hand in?
 - **Option A:** 10 page report + mockup (single column, single line spacing, 12 point font)
 - **Option B:** 5 page report + code and screenshots of the web application in a zipped folder
 - Submission on T-Square

Assignment II

- **Due: November 16, 2016**
- Questions will focus on the topics of the reading materials covered until the point the assignment is released.
- Data analytic questions – **Option A:** One question that would test your thinking and formulation of applying data analytic approaches to a social computing question
- Data analytic questions – **Option B:** Implement data analysis on a small dataset that will be provided

- What to hand in?
 - **Option A:** 10 page report (single column, single line spacing, 12 point font)
 - **Option B:** 5 page report + code and charts etc. in a zipped folder
 - Submission on T-Square

Term Project

- Project proposal due: September 21, 2016
- Mid-term presentations due: October 31, 2016
- Mid-term report due: October 31, 2016
- Final presentations due: December 12, 2016
- Final project report due: December 12, 2016
- Goals:
 - Group project: 2-4 people
 - You are free to pick your group – use Piazza to know and find like-interested classmates
 - Need to discuss your project idea with instructor early on in the course
 - Presentations (both midterm and final) will need to focus on the project goals, challenges faced, and task accomplished; a few additional minutes for Q&A

Term Project

- Individual assessment – clearly articulate individual goals and contributions in the project proposal
 - In midterm and final reports, revisit the above list to indicate what you have done
- Peer assessment – indicate how each of your groupmate accomplished what they were supposed to do
 - Only required in the final report

Finding Data

- Based on the project, you may need data.
- Suggestions:
 - Your own social media data. E.g., you might consider downloading all of your email or IM logs. Or, you could download all your Facebook data.
- Hillary Mason's research dataset page: <https://bitly.com/bundles/hmason/1>
- Jure Leskovec's SNAP page: <https://snap.stanford.edu/data/>

Late Policy

- Reading responses are due *at 11:59pm* on the day before the relevant class meeting.
- Assignments are due *at 11:59pm* on the date listed for that assignment.
- Term project reports (midterm, final) are due at 11:59pm on the date listed.
- Some work (only assignments, or the project deliverables, but not the presentations) submitted more than 15 minutes after the due time will be assessed a 25% penalty.
 - Each additional 24 hours of lateness will result in an additional 25% being taken off the grade for that assignment. After 2 days, the assignment will not be accepted and a grade of 0% will be entered.
- No extensions for reading reflections.

English as Second Language

- If English is not your first language, you may request to not be graded on your writing for a particular individual assignment.
 - This means you won't be penalized for bad writing, but you also won't get credit for good writing. To take advantage of this option, you must mark "ESL" (English as a Second Language) on the first page of your assignment/paper.
- This option is not available for the term project as it is a group assignment.

Academic Integrity

- This class abides by the Georgia Tech Honor Code.
- All assigned work is expected to be individual, except where explicitly indicated otherwise.
- You are encouraged to discuss the assignments with your classmates; however, what you hand in should be your own work.
 - Okay to use open-source software (no need to reinvent the wheel), however do acknowledge!
 - Copying/reusing code from your classmates and friends are not allowed; strict action will be taken if similarities are discovered
 - Copying (textual) content for your assignments and project from other published work (without citing them) is also not allowed, and is considered plagiarism

Help and Resources

- Office hours: By appointment
- Location:TSRB 341

- Teaching Assistant: Stevie Chancellor
- Office hours: By appointment
- Location:TSRB 341A
- Email: schancellor3@gatech.edu

- Email announcements will be made over the course page on T-Square

Part II: Introductions

name + program

academic background

closest social computing project, if any

what you want to learn from the class

Part II: Defining “Social Computing”

Quoting Wikipedia:

“Social computing is an area of computer science that is concerned with the intersection of social behavior and computational systems. It is based on creating or recreating social conventions and social contexts through the use of software and technology.”

A) An example of a social computing system (that exists online)

B) An example of a non-social computing system (that exists online)

Topics to be covered

- Key attributes/characteristics
 - Social networks and social media overview; Design of social systems
 - Public displays; Social capital; Identity, deception and performance
- Methods
 - Statistics, data mining review
 - Text analysis; network analysis
- Key issues of social computing
 - Credibility; trust; polarization; privacy
- Uses of social computing systems
 - Personality, emotion and behaviors; trends and forecasting; event analytics; social news; location and mobility; collaboration, volunteerism and activism

Suggested books

- SNA:
 - *Networks, Crowds, and Markets*, by David Easley and Jon Kleinberg
 - *Six Degrees*, by Duncan Watts
- Social Science-y:
 - *On Individuality and Social Forms*, by Georg Simmel
 - *Networked*, by Barry Wellman
- Machine Learning:
 - *Machine Learning for Hackers*, by Drew Conway and John Myles White
 - *Natural Language Processing with Python*, by Steven Bird, Ewan Klein, and Edward Loper
 - *Pattern Classification*, by Richard Duda, Peter Hart, and David Stork
- Pop Statistics:
 - *The Signal and the Noise*, by Nate Silver
- Writing:
 - *Writing for Social Scientists*, by Howard Becker

Handout – a survey that you need to fill out
and bring in the next class