AIT 722: ***Theories and Models in Geo-Social Data Analytics***

Instructor: Prof. Myeong Lee
Instructor E-mail: mlee89@gmu.edu
Instructor Office: ENGR 5304
Instructor Office Hours: By Appointment
TA: Rafael Chen (hchen28@gmu.edu)

Meeting Day/Time: online synchronous class

- Online synchronous sessions on Mondays from 4:30 PM – 7:10 PM (will be also recorded)

Meeting Method:

- Online sessions: Instructor will send out calendar invites with Zoom links

Catalog Description

This course introduces a broad spectrum of theories, conceptual models, machine learning, and computational modeling that are used in and related to geo-social data. Course contents include discussions of, and hands-on exercise with, geo-social data analytics, map-based visualization, community dynamics models, smart cities theories, and GIS-based system development. This course aims to help students grow as IT professionals who can (1) understand critical issues in smart and connected communities (S&CC), (2) combine data-driven approaches in understanding and addressing S&CC problems, and (3) communicate the geographically-embedded social patterns based on data analysis results through visualizations and interactive systems.

Prerequisites

Although there are no particular prerequisites for this course, basic statistical and data analysis knowledge is needed for understanding the course materials.

- *Prerequisite(s): Graduate Standing*
- *Recommended previous courses: Programming (Python or R); Statistics*

Rationale and Objectives

This course introduces a broad spectrum of theories, conceptual models, methodologies, computational modeling, and data analysis methods that are used in and related to geospatial/social data, particularly in the smart and connected communities (S&CC) domain. National Science Foundation (NSF) defines a smart and connected community as "a community that synergistically integrates intelligent technologies with the natural and built environments, including infrastructures, to improve the social, economic, and

environmental well-being of those who live, work, or travel within it" (NSF, 2019). Accordingly, this course covers geo-social data analytical topics as part of the S&CC curriculum in the field of Information Sciences and Technology. Mainly, this course is designed for IT professionals who are at the intersection between technology, people, and organizational management. Through this course, graduate students will have opportunities to learn about theoretical models and computational methods that together help understand the dynamics of S&CC and the characteristics of geo-tagged social data. Also, students will have opportunities to work on hands-on projects that integrate analysis results with map-based data visualization or interactive information systems. The final goal of this course is to help students grow as information professionals who can (1) understand and identify critical issues in geo-social data and the S&CC domains, (2) combine data-driven approaches in understanding and addressing the problems, and (3) communicate the issues and analysis results through map-based visualizations and interactive systems.

Topic Spectrum of the Course

The course includes but is not limited to the following topics and covers theoretical and technical aspects of geo-social- and community-oriented projects:

- Concepts of smart cities, local communities, and urban data science
- Computational modeling of urban/community dynamics
- Statistical methods for geo-social data analysis
- Geospatial data processing using GIS and programming languages
- Models and theories about local communities and community characteristics
- Map-based visualization for urban/community dynamics

Course assignments include:

- Reading and summarizing book chapters or research papers on urban/geo-social data topics
- Discussions on urban/community/geo-social data topics
- Geo-social data analysis and modeling
- Conceptualizing urban/community characteristics
- S&CC-related projects (either research or practical)
- Geospatial visualization using web- or desktop-based technologies

Course Format

This course is designed as a combination of lecture, student seminar, a technology-focused session, where the instructor provides an overview of the topics in each session (about 0.9 hours), a student team present to and discuss one paper/article from the reading list with the class (about 0.6 hours), and a technology-focused, hands-on session (about 1 hour). Each team of students (2-3 students/team) will start working on thinking about the final project at the beginning of the semester, and final deliverable will be a S&CC-related project website that combines theoretical, data analytical, or system developmental work. There will be about 16 sessions (2.5 hours per week) during the semester,

including mid-term/final presentations and/or guest lectures. Also, there will be four to five individual assignments.

Class Participation

Students are required to participate in course sessions and actively participate in the discussions. Also, students are encouraged to participate in asynchronous discussions through Blackboard.

Expectations

- Technological Expectations
  - Some familiarity with statistics, data management, databases, and Linux systems.
  - Programming concepts (Python or R)
  - Self-learning for unfamiliar technologies
- General Expectations
  - Each session will be provided with the assumption that you have read papers.
  - Lecture slides from instructor's material will be posted on Blackboard.
  - E-mail the instructor if you anticipate being unable to meet any course requirements in a timely manner.
- Safety and Security
  - Personal Safety and Security: The Mason Alert system provides emergency information of various sorts. Students can sign up for it by visiting the website https://alert.gmu.edu. Students are also reminded that an emergency poster exists in each classroom explaining what to do in the event of crises and that further information about emergency procedures exists on https://ready.gmu.edu/be-prepared/
  - Computer and IT Security: Visit GMU's IT http://itsecurity.gmu.edu/ web site regularly. Norton AntiVirus Software is free to download for all GMU students/faculty/staff.

Grading Policy

Student grades will be determined based on class participation, class presentation, assignments, and mid-term/final project presentations/reports.

Grade Component	Weight
General class participation and discussions (online/offline)	10%
Individual assignments (4-5 times)	20%
Team-based class presentations (seminars in regular sessions)	25%
Mid-term presentation & report (project proposal)	20%
Final presentation & report (project outcome)	25%

Grading Guidelines

Some grade components could be evaluated *subjectively.*

- A: consistently above and beyond the course/assignment requirements
- B: meets and occasionally exceeds the course/assignment requirements
- C: minimally meets the course/assignment requirements
- F: fails to meet the course/assignment requirements

* Grades will be awarded in accordance with the Mason Grading System for graduate students. See the university catalog for policies: http://catalog.gmu.edu for more information.

Grading Scale/Schema

The grading scale for this course is:

Grade Distribution		
97 – 100%	A+	Passing
94 – 96.99%	A	Passing
90 – 93.99%	A-	Passing
87 – 89.99%	B+	Passing
83 – 86.99%	B	Passing
77 – 82.99%	B-	Passing
70 – 76.99%	C	Passing
0 – 69.99%	F	Failing

Raw scores may be adjusted by the Instructor to calculate final grades. Students are responsible for checking the currency of their grade books. Grade discrepancies must be brought to instructor's attention within one week of assignment submission and 48 hours of exam submission.

Course Schedule and Contents

* The reading list is subject to change slightly over time as new studies are published.

Session	Topics	Content Description	Technology Covered
1	Overview and Course Logistics	• BlackBoard • Reading List ○ Kling, R., (2007) "What is Social Informatics", *The Information Society: An International Journal," Vol. 23* Iss: 4, pp. 205 – 220 ○ Foth, M., Choi, J. H. J., & Satchell, C. (2011, March). Urban informatics. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 1-8). ACM. ○ De Moor, A. (2009). Moving community informatics research forward. *The Journal of Community Informatics*, *5*(1). ○ Zheng, Y., Capra, L., Wolfson, O., & Yang, H. (2014). Urban computing: concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *5*(3), 38. • Class Activities	Git (using command line)

		○ Team Composition for projects as well as in-class seminars ○ Git process for team-based work	
2	Concepts of S&CC	• Reading List ○ Green, Ben. (2019). "1. The Smart City: A New Era on the Horizon." In *The Smart Enough City*. MIT Press. ○ Kitchin, Rob. 2016. "The Ethics of Smart Cities and Urban Science." *Philosophical Transactions A 374*. rsta.royalsocietypublishing.org. ○ Mohanty, Saraju P., Uma Choppali, and Elias Kougianos. 2016. "Everything You Wanted to Know about Smart Cities: The Internet of Things Is the Backbone." *IEEE Consumer Electronics Magazine 5*(3): 60–70.	APIs for S&CC Data (R, Python, or Websites) Introduction to Jekyll: Hello World
3	Computational Modeling of Cultural Dimensions in Communities	• Reading List ○ Christopher A Bail. (2014). The cultural environment: Measuring culture with big data. *Theory and Society, 43*(3-4):465-482. ○ Minsu Park, Ingmar Weber, Mor Naaman, and Sarah Vieweg (2015). Understanding musical diversity via online social media. In *ICWSM*. AAAI. ○ Minsu Park, Jaram Park, Young Min Baek, and Michael Macy. (2017). Cultural values and cross-cultural video consumption on Youtube. *PLoS one, 12*(5):e0177865. ○ Desislava Hristova, Luca M. Aiello, and Daniele Quercia. (2018). The new urban success: How culture pays. *Frontiers in Physics, 6*. doi: 10.3389/fphy.2018.00027.	Computational models to quantify cultural dimensions using Meetup data (R) Introduction to Jekyll: Data management
4	Prediction of Community Characteristics	• Reading List ○ Joshua Blumenstock, Gabriel Cadamuro, and Robert On. Predicting poverty and wealth from mobile phone metadata. *Science, 350*(6264):1073-1076. ○ Daniele Quercia, Joao Paulo Pesce, Virgilio Almeida, and Jon Crowcroft. (2013). Psychological maps 2.0: a web engagement	Machine Learning to predict socio-economic status (R) Jekyll-based Web Styling

		enterprise starting in London. In *Proceedings of the 22nd International Conference on World Wide Web*, 1065-1076. ○ Victor Soto, Vanessa Frias-Martinez, Jesus Virseda, and Enrique Frias-Martinez. (2011). Prediction of socioeconomic levels using cell phone records. *International Conference on User Modeling, Adaptation, and Personalization*, 377-388.	and Themes
5	Identifying Hidden Patterns of Communities	• Reading List ○ Justin Cranshaw, Raz Schwartz, Jason I Hong, and Norman Sadeh. (2012). The livehoods project: Utilizing social media to understand the dynamics of a city. *International AAAI Conference on Web and Social Media*. ○ Grant McKenzie, Zheng Liu, Yingjie Hu, and Myeong Lee. (2018). Identifying urban neighborhood names through user-contributed online property listings. *ISPRS International Journal of Geo-Information, 7*(10):388. ○ Anastasios Noulas, Salvatore Scellato, Renaud Lambiotte, Massimiliano Pontil, and Cecilia Mascolo. *A tale of many cities: universal patterns in human urban mobility. PLOS ONE, 7*(5):e37027, 2012.	Clustering analysis to find dynamic neighborhood (R) Jekyll-based Data Visualization
6	Theoretical Lens for S&CC	• Reading List ○ Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S., & Rickne, A. (2008). Analyzing the functional dynamics of technological innovation systems: A scheme of analysis. *Research policy, 37*(3), 407-429. ○ Le Dantec, C. (2012, May). Participation and publics: supporting community engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1351-1360). ○ Jenkins, T., Le Dantec, C. A., Disalvo, C., Lodato, T., & Asad, M. (2016, May). Object-oriented publics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing*	Low-Fi design of technology (Wireframe) Jekyll-based Navigation Techniques

		Systems (pp. 827-839). ACM. ○ Bowker, G. C., Baker, K., Millerand, F., & Ribes, D. (2009). Toward information infrastructure studies: Ways of knowing in a networked environment. In *International handbook of internet research* (pp. 97-117). Springer, Dordrecht.	
7	**Mid-term Presentation**	• Each team presents a proposal for the final project that aims to resolve S&CC-related problems (either practical or scientific). • As part of the proposal, each team needs to present a low-fidelity layout of how the final webpage would look like.	
8	Community-based System Design	• Reading List ○ John M Carroll, Michael Horning, Blaine Hoffman, Craig Ganoe, Harold Robinson, and Mary Beth Rosson. (2011). Community network 2.0: visions, participation, and engagement in new information infrastructures. *International Symposium on End User Development*, pages 270-275. ○ Yuheng Hu, Shelly D Farnham, and Andres Monroy-Hernandez. (2013). Whoo.ly: facilitating information seeking for hyperlocal communities using social media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3481-3490. ○ Asad, M., Le Dantec, C. A., Nielsen, B., & Diedrick, K. (2017, May). Creating a sociotechnical API: Designing city-scale community engagement. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 2295-2306). ACM.	Twitter data filtering techniques (R) Jekyll-based Graph Visualization
9	Place & Space	• Reading List ○ Dourish, P. (2006, November). Re-space-ing place: place and space ten years on. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work* (pp. 299-308). ACM. ○ Harrison, S., & Tatar, D. (2008). Places:	Modeling "placeness" using Meetup data (R) Jekyll-based Blog Posting

		people, events, loci–the relation of semantic frames in the construction of place. In *Journal of Computer Supported Cooperative Work (CSCW), 17*(2-3), 97-133. ○ Blaschke, T., Merschdorf, H., Cabrera-Barona, P., Gao, S., Papadakis, E., & Kovacs-Györi, A. (2018). Place versus space: From points, lines and polygons in GIS to place-based representations reflecting language and culture. *ISPRS International Journal of Geo-Information, 7*(11), 452. ○ Le Falher, G., Gionis, A., & Mathioudakis, M. (2015, April). Where is the Soho of Rome? Measures and algorithms for finding similar neighborhoods in cities. In *Ninth International AAAI Conference on Web and Social Media (ICSWM)*.	
10	Visualizing and Identifying Placeness	• Reading List ○ McKenzie, G., Janowicz, K., Gao, S., Yang, J. A., & Hu, Y. (2015). POI pulse: A multi-granular, semantic signature–based information observatory for the interactive visualization of big geosocial data. Cartographica: *The International Journal for Geographic Information and Geovisualization, 50*(2), 71-85. ○ Gorko, T., Yau, C., Malik, A., Harris, M., Tee, J. X., Maciejewski, R., ... & Ebert, D. (2018, January). A Multi-Scale Correlative Approach for Crowd-Sourced Multi-Variate Spatiotemporal Data. In *Proceedings of the 51st Hawaii International Conference on System Sciences*. ○ Yuan, J., Zheng, Y., & Xie, X. (2012, August). Discovering regions of different functions in a city using human mobility and POIs. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 186-194). ACM.	Map-based visualization using jQuery and Leaflet Jekyll-based Map Visualization

11	Bias in Geo- & Community-based Data	• Reading List ○ Hecht, B., & Stephens, M. (2014). A Tale of Cities: Urban Biases in Volunteered Geographic Information. In *Eighth International AAAI Conference on Weblogs and Social Media (ICWSM)*. Retrieved from http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8114 ○ Das, M., Hecht, B., & Gergle, D. (2019, April). The Gendered Geography of Contributions to OpenStreetMap: Complexities in Self-Focus Bias. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (p. 563). ACM. ○ Wang, G., Schoenebeck, S. Y., Zheng, H., & Zhao, B. Y. (2016, March). "Will Check-in for Badges": Understanding Bias and Misbehavior on Location-Based Social Networks. In *Tenth International AAAI Conference on Web and Social Media (ICWSM)*.	Calculating bias in the data using Census data and social media data (R) Jekyll-based Interactive Map Visualization
12	Community & Civic Engagement in S&CC	• Reading List ○ Gordon, E., Baldwin-Philippi, J., & Balestra, M. (2013). Why we engage: How theories of human behavior contribute to our understanding of civic engagement in a digital era. *Berkman Center Research Publication*, (21). ○ Patrick Sturgis, Ian Brunton-Smith, Jouni Kuha, and Jonathan Jackson. Ethnic diversity, segregation and the social cohesion of neighbourhoods in London. *Ethnic and Racial Studies, 37*(8):1286– 1309, 2014. doi: 10.1080/01419870.2013.831932 ○ Kavanaugh, A., Carroll, J. M., Rosson, M. B., Zin, T. T., & Reese, D. D. (2005). Community networks: Where offline communities meet online. *Journal of Computer-Mediated Communication, 10*(4), JCMC10417. ○ Mamei, M., Pancotto, F., De Nadai, M., Lepri, B., Vescovi, M., Zambonelli, F., & Pentland,	Running regressions for civic engagement (R) Components of Project-based Websites (examples)

		A. (2018). Is social capital associated with synchronization in human communication? An analysis of Italian call records and measures of civic engagement. *EPJ Data Science, 7*(1), 25.	
13	Community Diversity, Social Trust, and Social Capital	• Reading List ○ Dinesen, P. T., & Sønderskov, K. M. (2018). Ethnic diversity and social trust: a critical review of the literature and suggestions for a research agenda. *The Oxford handbook on social and political trust*, 175-204. ○ Gereke, J., Schaub, M., & Baldassarri, D. (2018). Ethnic diversity, poverty and social trust in Germany: Evidence from a behavioral measure of trust. *PloS ONE, 13*(7), e0199834. ○ Alessandri, G., Filosa, L., Tisak, M. S., Crocetti, E., Crea, G., & Avanzi, L. (2020). Moral disengagement and generalized social trust as mediators and moderators of rule-respecting behaviors during the COVID-19 outbreak. *Frontiers in psychology, 11*, 2020.	Network analysis to quantify social capital (R) Network visualization on Jekyll
14	Mobilities in and across S&CC	• Reading List ○ Chang, S., Pierson, E., Koh, P. W., Gerardin, J., Redbird, B., Grusky, D., & Leskovec, J. (2021). Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, *589*(7840), 82-87. ○ Zagheni, E., Garimella, V. R. K., & Weber, I. (2014, April). Inferring international and internal migration patterns from twitter data. In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 439-444). ACM. ○ Sheller, M., & Urry, J. (2006). The new mobilities paradigm. *Environment and planning A, 38(*2), 207-226. ○ Noulas, A., Shaw, B., Lambiotte, R., & Mascolo, C. (2015, May). Topological properties and temporal dynamics of place networks in urban environments. In *Proceedings of the 24th International*	Geospatial data analysis to understand mobility (R)

		Conference on World Wide Web (pp. 431-441). ACM.	
15	**Final Presentation**	• Each team presents the final project. • Final presentation needs to be comprehensive and justifiable. Also, it needs to provide appropriate analytics and solutions to answer or address particular questions in the local communities and S&CC domain. • The final website running on Github needs to be appealing and understandable to stakeholders (or general public) in the way that resolves the proposed problem.	

Academic Integrity and Honor Code

Mason is an Honor Code university; please see the Office for Academic Integrity (GMU's Honor Code) for a full description of the code and the honor committee process. The principle of academic integrity is taken very seriously and violations are treated gravely. What does academic integrity mean in this course? Essentially this: when you are responsible for a task, you will perform that task. When you rely on someone else's work in an aspect of the performance of that task, you will give full credit in the proper, accepted form. Another aspect of academic integrity is the free play of ideas. Vigorous discussion and debate are encouraged in this course, with the firm expectation that all aspects of the class will be conducted with civility and respect for differing ideas, perspectives, and traditions. When in doubt (of any kind) please ask for guidance and clarification.

Students are expected to do their own work in the course unless a group project is approved by the instructor. In papers and project reports, students are expected to write in their own words, rather than cutting-and-pasting from sources found on the Internet. The goal of assignments is to demonstrate what you have learned, not what you can google. When you do use text or graphical material from books, articles, and the Web, enclose the material in quotes and provide a complete and proper reference (in APA format). If a paragraph is used then it should be indented in the text (both left and right margins). In-text citation can use the [Author, Year] format or the Numerical [1] format which must refer to the source in the References section of your assignment. Use APA for guidance on citation style, usage, etc. (Don't buy the big CMS. See the smaller A Manual for Writers by Kate Turabian). Regardless of the citation method used, proper citations always include: Author(s), Title, Publication Date, Publisher, and URL (if from the Web, along with Last Accessed Date). BlackBoard's SafeAssign service will be used to review selected student assignments. The followings are additional honor code items:

- Wikipedia is **not** a primary reference. Use it for initial discovery, but use and cite primary references (which Wikipedia itself might use).
- If you need assistance with writing an assignment, you can get assistance here: http://writingcenter.gmu.edu
- Refer the Graduate Policies for general policies about courses and degrees: https://catalog.gmu.edu/policies/academic/graduate-policies/

- Any programming/coding assignments must adhere to the CS Honor Code.

Campus Closure

If the campus closes or class is canceled due to weather or other concern, students should check Blackboard [or other instruction as appropriate] for updates on how to continue learning and information about any changes to events or assignments.

Basic Course Technology Requirements (Online Courses)

Activities and assignments in this course will regularly use web-conferencing software (Blackboard Collaborate / Zoom). In addition to the requirements above, students are required to have a device with a functional camera and microphone. In an emergency, students can connect through a telephone call, but video connection is the expected norm.

Important Dates

Dates for adding, dropping the course, etc. are available via: https://registrar.gmu.edu

Religious Holidays

A list of religious holidays is available on the GMU's Religious Holiday Calendar. GMU respects any religious holidays. However, any student whose religious observance conflicts with a scheduled course activity must contact the Instructor at least 2 modules in advance of the conflict date in order to make alternative arrangements.

Course Materials and Student Privacy

All course materials posted to Blackboard or other course site are private to this class; by federal law, any materials that identify specific students (via their name, voice, or image) must not be shared with anyone not enrolled in this class.

- Videorecordings — whether made by instructors or students — of class meetings that include audio, visual, or textual information from other students are private and must not be shared outside the class
- Live video conference meetings (e.g. Collaborate or Zoom) that include audio, textual, or visual information from other students must be viewed privately and not shared with others in your household or recorded and shared outside the class

There is no textbook for this course, but research papers and book chapters will be used for the course materials. The reading list is presented in the weekly schedule.

The Instructor will not discuss issues relating to an individual student with anyone lacking a need to know without prior written permission of the student. This includes a student's family members and other students. Under no circumstances will a student's graded work be returned to another student.

Instructors, staff, and Teaching Assistants will take care to protect the privacy of each student's scores and grades.

Attendance Policy

Scheduled course sessions will be spent on clarification, amplification, and review of material through the use of slides, examples, and exercises. Lecture slides are complements to the lecture session, not substitutes for it. Each course session is an excellent time for you to raise questions, request additional examples, and get explanations of any ideas that are still unclear to you. As members of the academic community, all students are expected to contribute regardless of their proficiency with the subject matter. Students are expected to make prior arrangements with Instructor if they know in advance that they will miss any class and to consult with the Instructor as soon as possible if they miss any class without prior notice. Any student who expects to miss more than one class session is advised to drop the course and take it in a later semester when he/she can attend every class.

Disability Accommodations

Disability Services at George Mason University is committed to upholding the letter and spirit of the laws that ensure equal treatment of people with disabilities. Under the administration of University Life, Disability Services implements and coordinates reasonable accommodations and disability-related services that afford equal access to university programs and activities. Students can begin the registration process with Disability Services at any time during their enrollment at George Mason University. If you are seeking accommodations, please visit http://ds.gmu.edu/ for detailed information about the Disability Services registration process. Disability Services is located in Student Union Building I (SUB I), Suite 2500. Email:ods@gmu.edu | Phone: (703) 993-2474

Writing Center

A114 Robinson Hall; (703) 993-1200; http://writingcenter.gmu.eeu

Counseling and Psychological Services (CAPS)

(703) 993-2380 http://caps.gmu.eeu

References

National Science Foundation (2019). Smart & Connected Communities (S&CC) Program Solicitation. Retrieved from https://www.nsf.gov/pubs/2019/nsf19564/nsf19564.pdf