USC Viterbi
School of Engineering

DSCI 550: Data Science at Scale

Term-Day-Time
Summer 2022- Lecture: Wednesday and Friday, 2:00-6:10pm

Instructor: Dr. Anna Farzindar
Email: farzinda@usc.edu

Location: OHE100B and Virtual meeting (Please check the link on D2L)

Professor's Office Hours: Wednesday Before class
Office: virtual address
Students are advised to make appointments with the professor ahead of time in any event and be specific with the subject matter to be discussed. Students should also be prepared for their appointment by bringing all applicable materials and information.

DSCI 550 Overview

This course is designed as an overview course to give students a broad understanding of Informatics topics for Big Data and to get practical experience with key Big Data informatics techniques. Topics include roadmap of informatics, the data lifecycle, the role of the data scientist, and analyzing and exploring Big Data with real world use cases in data analytics, and big data. Understanding Big Data involves understanding of digital file formats, their detection and data extraction from them. Emphasis areas include Document Type Detection; Parsing and extraction; Metadata understanding and analysis; Language Identification and detection from files and finally file formats and representation. The class also has a specific focus on Content Detection and Analysis from large data sets. Datasets used in the course are publicly collected by the instructor or his collaborators involved in national Big Data initiatives including DARPA, NASA and

other projects. The course is designed to be accessible to students with experience programming in Python and Java at an intermediate level. The course will introduce the students to topical software frameworks that deal with Big Data including Tika, Solr, ElasticSearch™, TensorFlow, Nutch and Apache Hadoop™. The course will be a combination of lecture, in-class discussion, readings, group-based assignments and a final exam.

The objective of this course is to train students to be able to understand Big Data and Large Data Environments, e.g., file formats, their representation, and how to automatically extract information from large datasets of files. Specifically, students successfully completing this course will achieve three main objectives:

1. Develop sufficient proficiency in Big Data frameworks to write software capable of automatically extracting information from data including its text and metadata and language.
2. Develop sufficient proficiency in techniques with Large Data sets collected from the Web and other places (Intranet, Science Data Sets, Public Data Sets).
3. Develop sufficient proficiency in Python and Java to write and execute software that is “File Aware” and that automatically extracts text and metadata from large data sets.

The primary teaching methods will be discussion, case studies, and lectures. Students are expected to perform directed self learning outside of class which encompasses, among other things, a considerable amount of literature review. Leadership training in open source is provided and encouraged, and students leave with an experience in open source that makes them more marketable to companies and institutions looking to hire in Big Data, and Data Science.

In addition to foundations, and practical experience with Big Data and Data Science, the class will also introduce the student to the state-of-the-art in content detection research, future trends and state-of-the-practice. Students are expected to attend class regularly, and participate (as directed) in all class discussions, and most importantly, have fun!

Textbooks

Chris A. Mattmann, and Jukka Zitting. Tika in Action, 256 pages. New York: Manning Publications, November 2011. ISBN: 9781935182856.

C. Mattmann. Machine Learning with TensorFlow: 2nd Edition. 456 pages. New York: Manning Publications, December 2020. ISBN 9781617297717

ASSIGNMENTS And EXAMINATIONS

Name	Description	Weight
Exam	An exam testing your understanding of the lecture materials	25%
Assignments	Assignments where you will build on the Big Data and Data Science topics in course and make a contribution to one of the existing open source frameworks (Tika, Solr, Nutch, TensorFlow, OODT, etc.).	45%
Individual Presentation	An individual presentation demonstrating the student's understanding of one of the required paper readings in the course.	15%
Quizzes	Participation in lectures, by asking questions and contributing to the conversation. Attending lectures (physically and remotely) and positively contributing to the class experience.	15%

Schedule

(subject to change; check regularly)

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
Week1	• **Course Introduction** • **Introduction to Big Data** • Breakout Groups on Big Data	--	Resources: • DARPA I2O (DARPA Dan) Video • Square Kilometre Array Video • Kenneth Cukier: Big data is better data
	• Report out from Big Data Breakouts • **A Taxonomy of File Formats**	• Tika in Action, Chapter 1 • Mattmann, Chris. A vision for data science. Nature, Vol. 493, No. 7433, pp.	•

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
	• Content Detection Libraries • Language Bindings for Apache Tika • Individual Student Presentations • Report outs from the in class discussion around 5 ‘Vs’	473-475, January 24, 2013. • Lynch, Clifford. "Big data: How do your data grow?." Nature 455.7209 (2008): 28-29.	
Week 2	• **Document Similarity and Deduplication** • Individual Presentations • Introduction to Assignment 1	• Tika in Action, Chapter 2 • Howe, Doug, et al. "Big data: The future of biocuration." Nature 455.7209 (2008): 47-50. • Wigan, Marcus R., and Roger Clarke. "Big data's big unintended consequences." Computer 46.6 (2013): 46-53. • Schwartz, J. A. N. A., et al. "Measuring the value of Big Data exploitation systems: Quantitative, non-subjective metrics with the user as a key component." Parsons Journal for Information Mapping 6 (2014): 1-12. • Sotera Defense Solutions. A Survey of Big Data Methods, Assessments, and Approaches. November 2012 • De Mauro, Andrea, Marco Greco, and Michele Grimaldi.	Resources: • The MIME guys: How two Internet gurus changed e-mail forever • IEEE Computer Interview with Nathaniel Bornstein

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		"What is big data? A consensual definition and a review of key research topics." AIP conference proceedings. Vol. 1644. No. 1. AIP, 2015. • Crocker, David. RFC 822 "Standard for the format of ARPA Internet text messages." (1982). • Freed, Ned and Nathaniel Borenstein. RFC 1341. MIME (Multipurpose Internet Mail Extensions). Mechanisms for Specifying and Describing the Format of Internet Message Bodies. June 1992. • Freed, Ned, and Nathaniel Borenstein. RFC 2045. Multipurpose internet mail extensions (MIME) part one: Format of internet message bodies. 1996.	
	• **Document Type Detection** • **Advanced File System Statistics and Understanding** • Individual Presentations • Report outs from the in class discussion around classifying files	• Tika in Action, Chapter 3 • Freed, Ned, and Nathaniel Borenstein. RFC 2046 Multipurpose internet mail extensions (MIME) part two: Media types, November, 1996. • Freed, Ned. RFC 2048 "Multipurpose internet mail extensions (MIME) part four: Registration	Resources: • The Economist - Digital Bit Rot • Bit Rot - on Digital Vellum (by Vint Cerf, Google) TEDx Talk

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
	and the MIME taxonomy	procedures." ISI (1996). • Hicks, Ben J., et al. "Organizing and managing personal electronic files: A mechanical engineer's perspective." ACM Transactions on Information Systems (TOIS) 26.4 (2008): 23. • Shim, Jungwon Roy. "Arium: Beyond the Desktop Metaphor: A new way of navigating, searching, and organizing personal digital data." Masters Thesus, Carnegie Mellon University (2012). • Crowder, Jerome, Jonathan Marion, and Michele Reilly. "File Naming in Digital Media Research: Examples from the Humanities and Social Sciences." Journal of Librarianship and Scholarly Communication 3.3 (2015). • Jackson, Andrew N. "Formats over time: Exploring UK web history." arXiv preprint arXiv:1210.1714 (2012). • Bik, Elisabeth M., Casadevall, Arturo, Fang, Ferrie C. The Prevalence of Inappropriate Image Duplication in Biomedical Research Publications.Manku, Gurmeet Singh,	

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		Arvind Jain, and Anish Das Sarma. "Detecting near-duplicates for web crawling." Proceedings of the 16th international conference on World Wide Web. ACM, 2007.	
Week 3	• **Content Extraction** • Individual Presentations	• Tika in Action, Chapter 4 • Henzinger, Monika. "Finding near-duplicate web pages: a large-scale evaluation of algorithms." Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2006. • Cooper, Matthew, Jonathan Foote, and Andreas Girgensohn. "Automatically organizing digital photographs using time and content." Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on. Vol. 3. IEEE, 2003. • Manber, Udi. "Finding similar files in a large file system." Usenix Winter. Vol. 94. 1994. • Chim, Hung, and Xiaotie Deng. "Efficient phrase-based document	Resources: • Introduction to Information Retrieval - Chris Mattmann

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		similarity for clustering." IEEE Transactions on Knowledge and Data Engineering 20.9 (2008): 1217-1229. • Amirani, Mehdi Chehel, Mohsen Toorani, and A. Beheshti. A new approach to content-based file type detection. Computers and Communications, 2008. ISCC 2008. IEEE Symposium on. IEEE, 2008. • McDaniel, Mason, and M. Hossain Heydari. Content based file type detection algorithms. System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on. IEEE, 2003. • Alamri, Nasser S., and William H. Allen. "A comparative study of file-type identification techniques." SoutheastCon 2015. IEEE, 2015. • Li, Wei-Jen, et al. "Fileprints: Identifying file types by n-gram analysis." Information Assurance Workshop, 2005. IAW'05. Proceedings from the Sixth Annual IEEE SMC. IEEE, 2005. • Shahi, Ashim. "Classifying the classifiers for file fragment	

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		classification." Masters Thesis, Universiteit van Amsterdam (2012).	
	• Last Week Tonight: Robocalls • RoboKiller Video • **Understanding Metadata** • Individual Presentations	• Tika in Action, Chapter 5 • Ahmed, Irfan, et al. "Fast file-type identification." Proceedings of the 2010 ACM Symposium on Applied Computing. ACM, 2010. • Pierris, Georgios, and Stilianos Vidalis. "Forensically classifying files using HSOM algorithms." Emerging Intelligent Data and Web Technologies (EIDWT), 2012 Third International Conference on. IEEE, 2012. • Harris, Ryan M. "Using artificial neural networks for forensic file type identification." Master's Thesis, Purdue University (2007). • Douceur, John R., and William J. Bolosky. A large-scale study of file-system contents. ACM SIGMETRICS Performance Evaluation Review 27.1 (1999): 59-70. • Kilicoglu, Halil, et al. "Semantic MEDLINE: a web application for managing the results of PubMed	Resources: • Last Week Tonight - John Oliver (HBO) - Robocalls • New App to stop Robocalls - RoboKiller

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		Searches." Proceedings of the third international symposium for semantic mining in biomedicine. Vol. 2008. 2008. • Kobayashi, Mei, and Koichi Takeda. "Information retrieval on the web." ACM Computing Surveys (CSUR) 32.2 (2000): 144-173. • Voorhees, Ellen M., and Donna Harman. "Overview of the sixth text retrieval conference (TREC-6)." Information Processing & Management 36.1 (2000): 3-35. • Arasu, Arvind, and Hector Garcia-Molina. Extracting structured data from web pages. Proceedings of the 2003 ACM SIGMOD international conference on Management of data. ACM, 2003.	
Week 4	• Video - Understanding Metadata TedX talk • **Information Clustering** • **Exam Review** • Example - Machine Learning with TensorFlow - Activity	• Tika in Action, Chapter 6 • Lewandowski, Dirk. "Web searching, search engines and Information Retrieval." Information Services & Use 25.3, 4 (2005): 137-147. • Weninger, Tim, William H. Hsu, and Jiawei Han. "CETR: content extraction via tag ratios."	Resources: • The Power of Metadata: Deepak Jagdish and Daniel Smilkov at TEDxCambridge 2013

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
	Clustering (Android)	Proceedings of the 19th international conference on World wide web. ACM, 2010. • Karpathy, Andrej, and Li Fei-Fei. "Deep visual-semantic alignments for generating image descriptions." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. • Gowda, Thamme, and Chris A. Mattmann. "Clustering Web Pages Based on Structure and Style Similarity (Application Paper)." Information Reuse and Integration (IRI), 2016 IEEE 17th International Conference on. IEEE, 2016. • Anquetil, Nicolas, and Timothy Lethbridge. File clustering using naming conventions for legacy systems. Proceedings of the 1997 conference of the Centre for Advanced Studies on Collaborative research. IBM Press, 1997. • Swierk, Edward, et al. "The Roma personal metadata service." Mobile Networks and Applications 7.5 (2002): 407-418. • Karypis, Michael Steinbach George, Vipin Kumar, and	

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		Michael Steinbach. "A comparison of document clustering techniques." KDD workshop on Text Mining. 2000. • Marchionini, Gary. "Exploratory search: from finding to understanding." Communications of the ACM 49.4 (2006): 41-46.	
	Exam (June 10, 2022)		
Week 5	• Assignment 2 • Video - Linguistic Forensics • **Language Identification** • Individual Presentations	• Tika in Action, Chapter 7 • Koehn, Philipp, et al. "Moses: Open source toolkit for statistical machine translation." Proceedings of the 45th annual meeting of the ACL on interactive poster and demonstration sessions. Association for Computational Linguistics, 2007. • Post, Matt, et al. "Joshua 5.0: Sparser, better, faster, server." Proceedings of the Eighth Workshop on Statistical Machine Translation. 2013. • Lins, Rafael Dueire, and Paulo Gonçalves. Automatic language identification of written texts. Proceedings of the 2004 ACM symposium on Applied computing. ACM, 2004. • Papineni, Kishore, et al. "BLEU: a method	Resources: • Forensic Linguistic Profiling & What Your Language Reveals About You | Harry Bradford | TEDxStoke • The Shallowness of Google Translate - The Atlantic - Douglas Hofstadter • Apache cTAKES • Apache UIMA • Apache OpenNLP • Stanford Core NER/NLP • NLTK

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		for automatic evaluation of machine translation." Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics, 2002. • Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate." arXiv preprint arXiv:1409.0473 (2014). • Tromp, Erik, and Mykola Pechenizkiy. "Graph-based n-gram language identification on short texts." Proc. 20th Machine Learning conference of Belgium and The Netherlands. 2011. • Lopez-Moreno, Ignacio, et al. "Automatic language identification using deep neural networks." Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on. IEEE, 2014. • Bertoldi, Nicola, et al. "MMT: New open source MT for the translation industry." Proceedings of The 20th Annual Conference of the	

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		European Association for Machine Translation (EAMT). 2017.	
	• Individual Presentations • **Discussion on Named Entity Recognition** • Hadoop Spark and Tika: Large Scale Content Detection and Analysis	• Tika in Action, Chapter 8 • Tjong Kim Sang, Erik F., and Fien De Meulder. "Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition." Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4. Association for Computational Linguistics, 2003. • Nadeau, David, and Satoshi Sekine. "A survey of named entity recognition and classification." Lingvisticae Investigationes 30.1 (2007): 3-26. • Ritter, Alan, Sam Clark, and Oren Etzioni. "Named entity recognition in tweets: an experimental study." Proceedings of the Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2011. • Mattmann, Chris A., and Madhav Sharan. "An automatic approach for	Resources: • Apache Distributed Release Audit Tool (DRAT) • DRAT Video

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		discovering and geocoding locations in domain-specific web data." Proceedings of the 2016 IEEE 17th International Conference on Information Reuse and Integration (IRI'16). 2016. • Khodak, Mikhail, Nikunj Saunshi, and Kiran Vodrahalli. "A Large Self-Annotated Corpus for Sarcasm." arXiv preprint arXiv:1704.05579 (2017). • Hutto, Clayton J., and Eric Gilbert. "Vader: A parsimonious rule-based model for sentiment analysis of social media text." Eighth international AAAI conference on weblogs and social media. 2014. • Geyer, Kelly, et al. "Named Entity Recognition in 140 Characters or Less." # Microposts. 2016.	
Week 6	• Readout - Named Entity Recognition Group Presentations • Open Source Content Detection Technologies • Individual Presentations	• Tika in Action, Chapter 9 • Dean, Jeffrey, and Sanjay Ghemawat. MapReduce: simplified data processing on large clusters. Communications of the ACM 51.1 (2008): 107-113. • Zaharia, Matei, et al. Spark: cluster computing with	Resources: • TedXTalk Analyzing and modeling complex and big data | Professor Maria Fasli | TEDxUniversityofEssex • Textract • Scrapy • Mattmann ApacheCon 2015 Content Talk

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		working sets.Proceedings of the 2nd USENIX conference on Hot topics in cloud computing. Vol. 10. 2010. • Elsayed, Tamer, Jimmy Lin, and Douglas W. Oard. "Pairwise document similarity in large collections with MapReduce." Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers. Association for Computational Linguistics, 2008. • M. Bernaschi, M. Cianfriglia, A. Di Marco, A. Sabellico, G. Me, G. Carbone, G. Totaro. Forensic Disk Image Indexing and Search in an HPC environment. IEEE International Conference on High Performance Computing & Simulation (HPCS), 2014. • Meusel, Robert, Peter Mika, and Roi Blanco. "Focused crawling for structured data." Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge	

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		Management. ACM, 2014. • Niu, Feng, et al. "DeepDive: Web-scale Knowledge-base Construction using Statistical Learning and Inference." VLDS 12 (2012): 25-28. • Mattmann, C. A., Oh, J. H., Palsulich, T., McGibbney, L. J., Gil, Y., & Ratnakar, V. (2015, November). DRAT: An Unobtrusive, Scalable Approach to Large Scale Software License Analysis. In Automated Software Engineering Workshop (ASEW), 2015 30th IEEE/ACM International Conference on (pp. 97-101). IEEE.	
	• Ted Talk (click link in resources) • Evaluating Content Detection • Walkthrough of Polar Deep Insights • Individual Presentations	• Tika in Action, Chapter 10 • Białecki, Andrzej, et al. "Apache lucene 4." SIGIR 2012 workshop on open source information retrieval. 2012. • Turtle, Howard, Yatish Hegde, and S. Rowe. "Yet another comparison of lucene and indri performance." SIGIR 2012 Workshop on Open Source Information Retrieval. 2012. • Bontcheva, Kalina, et al. "TwitIE: An Open-	Resources: • This is What happens when you reply to SPAM e-Mail. - James Veitch - Ted Talk • USC IRDS Polar Data Science website

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		Source Information Extraction Pipeline for Microblog Text." RANLP. 2013. • Cunningham, Hamish. "GATE, a general architecture for text engineering." Computers and the Humanities 36.2 (2002): 223-254. • Atserias, Jordi, et al. "FreeLing 1.3: Syntactic and semantic services in an open-source NLP library." Proceedings of LREC. Vol. 6. 2006. • Manning, Christopher D., et al. "The stanford corenlp natural language processing toolkit." ACL (System Demonstrations). 2014. • Savova, Guergana K., et al. "Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications." Journal of the American Medical Informatics Association 17.5 (2010): 507-513.	
Week 7	• Group Readouts on Evaluating Content Detection and Analysis • Lecture on NoSQL	• Tika in Action, Chapter 11 • Nowell, Lucy Terry, et al. "Visualizing search results: some alternatives to query-document similarity." Proceedings of the	Resources: • 2019 MIT Citi - 3 - Top Ten Big Data Blunders - Michael Stonebraker (Only watch first 15m)

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
	• Lecture on SciSpark • Individual Presentations	19th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 1996. • Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." Visual Languages, 1996. Proceedings., IEEE Symposium on. IEEE, 1996. • Gottron, Thomas. "Evaluating content extraction on HTML documents." Proceedings of the 2nd International Conference on Internet Technologies and Applications (ITA'07). 2007. • Leuski, Anton. "Evaluating document clustering for interactive information retrieval." Proceedings of the tenth international conference on Information and knowledge management. ACM, 2001. • Bailey, Peter, et al. "Evaluating search systems using result page context." Proceedings of the third symposium on Information interaction in context. ACM, 2010.	• Data Manipulation at Scale - NoSQL rebuttal

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
	• Video - Scientific Data: Water and Snow in the Western US • Searching Scientific Datasets • Scientific Data Processing (Airborne Snow Observatory) • Individual Presentations • Big Data with an Eye Towards the Future: Discussion	• Tika in Action, Chapter 12 - 14 • Palamuttam, Rahul, et al. "SciSpark: Applying in-memory distributed computing to weather event detection and tracking." Big Data (Big Data), 2015 IEEE International Conference on. IEEE, 2015. • Leavitt, Neal. "Will NoSQL databases live up to their promise?." Computer 43.2 (2010). • Stonebraker, Michael. "SQL databases v. NoSQL databases." Communications of the ACM 53.4 (2010): 10-11. • Stonebraker, Michael. "Stonebraker on NoSQL and enterprises." Communications of the ACM 54.8 (2011): 10-11. • Rafique, Ansar, et al. "On the performance impact of data access middleware for nosql data stores." IEEE Transactions on Cloud Computing (2015). • Moniruzzaman, A. B. M., and Syed Akhter Hossain. "Nosql database: New era of databases for big data analytics-classification, characteristics and comparison." arXiv preprint	Resources: • Thomas Painter of NASA JPL speaks at TEDxIS on a thought provoking take

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		arXiv:1307.0191 (2013).	
Week 8	• Individual Presentations	• C. Mattmann, D. Freeborn, D. Crichton, B. Foster, A. Hart, D. Woollard, S. Hardman, P. Ramirez, S. Kelly, A. Y. Chang, C. E. Miller. A Reusable Process Control System Framework for the Orbiting Carbon Observatory and NPP Sounder PEATE missions. In Proceedings of the 3rd IEEE Intl Conference on Space Mission Challenges for Information Technology (SMC-IT 2009), pp. 165-172, July 19 - 23, 2009. • Wilkinson, Mark D., et al. "The FAIR Guiding Principles for scientific data management and stewardship." Scientific data 3 (2016): 160018. • Buneman, Peter, et al. "Archiving scientific data." ACM Transactions on Database Systems (TODS) 29.1 (2004): 2-42. • Fox, Peter, and James Hendler. "Changing the equation on scientific data visualization." Science 331.6018 (2011): 705-708. • Plale, Beth, et al. "Active management	•

Week	Lecture Topic	Assigned Readings	Assignments & Deadlines
		of scientific data." IEEE Internet Computing 9.1 (2005): 27-34. • Gray, Jim, et al. "Scientific data management in the coming decade." ACM SIGMOD Record 34.4 (2005): 34-41. • Ailamaki, Anastasia, Verena Kantere, and Debabrata Dash. "Managing scientific data." Communications of the ACM 53.6 (2010): 68-78.	

USC ACADEMIC INTEGRITY

Statement on Academic Conduct and Support Systems

Academic Conduct Plagiarism - presenting someone else's ideas as your own, either verbatim or recast in your own words - is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Section 11, Behavior Violating University Standards. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct. Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the Office for Equity, Equal Opportunity, and Title IX or to the Department of Public Safety. This is important for the safety of the USC community. Another member of the university community - such as a friend, classmate, advisor, or faculty member - can help initiate the report, or can initiate the report on behalf of another person. The Sexual Violence Prevention & Services provides 24/7 confidential support, and the sexual assault resource center webpage sarc@usc.edu describes reporting options and other resources.

Support Systems

A number of USC's schools provide support for students who need help with scholarly writing. Check with your advisor or program staff to find out more. Students whose primary language is not English should check with the American Language Institute which sponsors courses and workshops specifically for international graduate students. The Office of Disability Services and Programs provides certification for students with disabilities and helps arrange the relevant accommodations. If an officially declared emergency makes travel to campus infeasible, USC Emergency Information will provide safety and other updates, including ways in which instruction will be continued by means of blackboard, teleconferencing, and other technology.

Statement on Diversity

The diversity of the participants in this course is a valuable source of ideas, problem solving strategies, and engineering creativity. We encourage and support the efforts of all of our students to contribute freely and enthusiastically. We are members of an academic community where it is our shared responsibility to cultivate a climate where all students and individuals are valued and where both they and their ideas are treated with respect, regardless of their differences, visible or invisible.