Course Title: Computational Tools for Cultural Comparison

Course Number: COREI-AD 52J

Course satisfies: Experimental Discovery in the Natural World

Location: NYU Abu Dhabi.

Credit hours: 4

Prerequisites: None

Co-requisites: None

Office hours: By appointment.

Name and Contact Information of Instructor:

 Prof. Dennis Shasha

 Department of Computer Science

 Courant Institute of Mathematical Sciences

 New York University

 251 Mercer Street

 New York, N.Y. 10012-1110

 U.S.A.

 Email: shasha@cs.nyu.edu

 Web: http://cs.nyu.edu/cs/faculty/shasha/index.html

Brief Course Description:

Different cultures have both unique and common properties. These can be linguistic properties, properties of politeness, of government, of art, of economics, of food, of gender relationships, and even of war. Tools for cross-cultural comparison should allow many individuals, each of whom is familiar with only a few cultures to enter properties and values about those cultures and then allow searchers to compare the cultures in various ways. For example, it should be possible to discover the commonalities and differences among pairs of cultures, to cluster many cultures based on their commonalities, and to find cultural rules of the form "if culture has property x, then it has property y".

This course will begin with a discussion of a key cultural notion -- language. You will learn the basics of linguistics and then will interact with a cross-cultural database and analysis platform called Terraling. We will also study case studies of the cross-cultural influence on management practices and the art of negotiation. After that, you will engage in two person projects to use Terraling to enter properties and values about a topic of your choice (e.g. about food, legal structure, health policy etc.). Some of that data will come from the data sets others have gathered that are on the web and some will come from data that you yourselves gather. In the process you will learn elements of statistics, machine learning, programming in python, and data science.

Learning Outcomes:

Learn to deal with multiple data sources to integrate them computationally.

Understand the basics of computational data science, particularly data organization, probability, statistics, and machine learning.

Achieve basic fluency as a programmer.

Teaching and Learning Methodologies:

Lectures, readings, data gathering, programming instructions.

Methods and Dates of Assessments:

Class and presentation work 20% (evaluated based on interaction with other presenters and quality of student’s own presentations), in-class quizzes 15%, week 1 project 10%, week 2 project 20%, and week 3 project 35%.

Information on out-of-class assignments: Daily readings and programming assignments.

Project 1: Identify the two or more data sources you will use for your final project, explain your final design in terms of culture, property, values. Explain your use of the examples feature of Terraling. Identify the programming knowledge you will need.

Project 2: Describe the design of the python programs that will import the data and transform it. Illustrate components of the code that can parse the various data sources to put it into a format that Terraling can use.

Project 3: Presentation of your system in Terraling, showing the queries done and insights gained from the analytical tools that are built in and any additional analytical tools you have built, especially statistical ones. A respondent team will discuss the queries they have tried and insights they might have gained.

Because of the writing requirement common to all core courses, accompanying each formal presentation of a project and each class presentation of a book chapter, you will also provide a three page description of what you are going to say, your sources, and so on. Your presentation may refer to that essay, but please don’t read from it word for word. The Global Academic Writing Fellows can help you with the written part of your work.

Course texts:

# 1. *The Atoms of Language: The Mind's Hidden Rules of Grammar*

Mark Baker

Paperback: 288 pages

Publisher: Basic Books; Reprint edition (October 8, 2002)

Language: English

ISBN-10: 0465005225

ISBN-13: 978-0465005222

2. *Cross-Cultural Management: Essential Concepts*

David Thomas and Mark Peterson

Paperback: 344 pages

Publisher: SAGE Publications, Inc; Third Edition edition (February 19, 2014)

Language: English

ISBN-10: 1452257507

ISBN-13: 978-1452257501

3. *Culture and Conflict Resolution*

Kevin Avruch

Paperback: 172 pages

Publisher: United States Institute of Peace (November 1998)

Language: English

ISBN-10: 1878379828

ISBN-13: 978-1878379825

4. *Puzzles for Programmers and Pros*

 Dennis Shasha

 Publisher: Wrox; 1 edition (May 7, 2007)

 Language: English

 ISBN-10: 0470121688

 ISBN-13: 978-0470121689

5. *Statistics is Easy*, second edition

Dennis Shasha and Manda Wilson

Publisher: Morgan & Claypool [can be obtained online]

Language: English

ISBN-10: 1598297775

ISBN-13: 978-0393336832

Course topics and Contents on a Week-by-Week Basis:

Week 1: Understanding commonalities and differences across languages. Gaining familiarity with the use and structure of Terraling (Syntactic Structures application) for cross-linguistic analysis. A first principes approach to probability. Begin to learn python through graduated examples.

Primary readings: 1. *Puzzles for Programmers and Pros* 2. The website http://anthropology.ua.edu/cultures/cultures.php?culture=Cross-Cultural%20Analysis 3. *Atoms of Language* 4. www.terraling.com

End-of-week project: Project 1 (see above)

Day 1: Introduction to cross-cultural analysis – questions asked, critiques. Fundamentals of linguistics. Terraling and the goal of cross-linguistic analysis. <http://anthropology.ua.edu/cultures/cultures.php?culture=Cross-Cultural%20Analysis> Install python, introduction to the shell, types and hello world.

Day 2: Student presentations of chapter 1 of *Atoms of Language*. Terraling as applied to syntactic analysis. Interaction between types, list and string methods, conditionals.

Day 3: Student presentations of chapters 2 and separately chapter 3 of *Atoms of Language*. Terraling syntactic analysisIntroduction to modules, how to build a cypher, loops.

Day 4: Student presentation of chapters 4 and separately of chapter 7 of *Atoms of Language*. Present project idea, data sources.

Python it all together (exercise day: fizzbuzz, Celsius & Farenheit). Python file treatment and parsing.

Week 2: Approaches to cross-cultural management. Applying statistics to cross-cultural phenomena. In parallel, we will teach more elements of the Python programming language including elements of parsing poorly formatted data.

Primary readings: 1. *Cross-Cultural Management: Essential Concepts* 3. *Statistics is Easy*

Day 1:.Basic notions of statistics using a resampling approach. Monte Hall Puzzle and how to conceive of probabilities. .Lucky Roulette. Python maps & functions.

Day 2: Student presentation of both chapters 1 and 2 of *Cross-Cultural Management*. More probability: bait and switch, feedback dividends. Python data integration. Chapters 1 and 2 of *Statistics is Easy*

Day 3: Student presentation of chapter 3 of *Cross-Cultural Management* Computational game-playing techniques. Python data integration.

Day 4: Student presentation of chapter 6 of *Cross-Cultural Management* . Python interactive games.

Day 5: Dependency theory in databases. Student presentation of chapter 7 of *Cross-Cultural Management.* Python exercise day.

End-of-week project: Project 2 (see above)

Week 3: Culturally sensitive approaches to negotiation in international affairs. Data handling for the final project. Presentation for the final project.

Primary reading: *Culture and Conflict Resolution*

Day 1: Student presentation of Part I of *Culture and Conflict Resolution*. Introduction to information theory and machine learning. Python data analytical tools.

Day 2: Student presentation of Part II of *Culture and Conflict Resolution* Python data analytical tools.

Day 3: Student presentation of Part III of *Culture and Conflict Resolution*. Python simulated annealing and knapsack.

Day 4: Project 3: Final student presentations along with the respondent team who used each system.

Crosslisting: relevant to economics and anthropology, and social science in general, but also to new literature.

Pedagogical methods to keep student interest: Course will be very interactive with lots of in-class time devoted to student presentations and projects.

Cultural diversity is explicitly addressed. The course builds on the cultural diversity of the students at NYU Abu Dhabi and asks them to research this aspect further.

Relationship of this course to others: close relationship to any course dealing with multiple cultures. Not specifically redundant to any other course.

Appendix

Data sources ICPSR : NYU is apparently a member but I'd have to find out how to get to the data. Inter-University Consortium For Political and Social Research <http://www.icpsr.umich.edu/icpsrweb/ICPSR/index.jsp>

Matthew Sumner matt.sumner@nyu.edu on Jan 3, 2015 said: “the students need only access ICPSR via the library site. Once there, it is possible to register for a free account. That's it.”

Demographic and Health Surveys http://www.measuredhs.com/start.cfm

Labor Statistics Data http://www.bls.gov/bls/other.htm

Personality tests: http://www.bbc.co.uk/science/humanbody/mind/index\_surveys.shtml

Correlates of war: http://www.correlatesofwar.org/

Population data across the world: http://sedac.ciesin.columbia.edu/data/collection/gpw-v3

Human genome data: http://hgdp.uchicago.edu/cgi-bin/gbrowse/HGDP/

Archaeological data: http://www.tdar.org/

United Nations statistical commission: <http://unstats.un.org/unsd/default.htm>

Cross cultural management journal articles e.g.: <http://www.academia.edu/1489279/CrossCultural_Management_in_practice_culture_and_negotiated_meanings?login=&email_was_taken=true>