New Graduate Course Proposal – Winter 2022

Data Visualization

Course Code: *BIOM 5403*

Course Title: *Data Visualization*

Type of Request: *New Course*

Calendar Style Description:

Fundamental concepts of information visualizations, including an introduction to visualization (importance, basic approaches, and existing tools), abstract visualization concepts, human perception, visualization methodology, interaction, design and development of interactive visualization systems, use of visualization software, tools, and systems. Use case studies will be drawn from various domains, including biomedical engineering, scientific, and business applications.

Pre-requisites:

Assumed pre-requisite knowledge: basic knowledge of programming.

Grading Scheme:

This course is research-based and focuses on active learning; thus, the evaluation is primarily focused on hands-on activities, presentation, analyzing/discussing research, and writing the report in a scientific form.

10% participation, 30% assignments, 60% presentations (in-class and video) & individual project.

- ***Participation and activities:*** *research paper discussion, in-class hands-on activities (e.g., ideating, sketching) (10%).*
- ***Assignments****: biomedical visualization research paper presentation (10%), critiques of visualizations (10%), literature review (visualization) (10%).*
- ***Presentations & Individual project:*** *student will apply their knowledge throughout the course to complete the project. The project has 5 deliverables throughout the term in order:*
    1. *project proposal (project idea, choice of dataset, reasoning) (10%),*
    2. *mid-term project presentation (conference presentation style) (10%),*
    3. *interactive visualization design and development (15%),*
    4. *final project presentation (conference presentation style) (10%),*
    5. *final written report (IEEE conference format) (15%).*

The midterm presentations will be done in class; each student has a 3-minute slot to present their work. The instructor will provide written feedback to the presenter on the content and their style of presentation. The final presentation will be in the form

of pre-recorded videos. Students will be asked to submit a 10–15-minute video of their project.

***All project ideas will be reviewed and approved by the instructor**. I will provide a list of **open-sourced biomedical datasets for student projects**. I anticipate students' proposed user studies to be minimal risk. I will apply for a "Course-Based Research Ethics" clearance.*

Text:

The following textbooks are strongly recommended:

- *Visualization Analysis and Design, Munzner, Tamara, A K Peters, Limited, 2014, 9781466508910*
- *Edward Tufte, Envisioning Information, Graphics Press, ISBN 0961392118*

The following useful references and resources will be used throughout the course:

- *Revisiting Bertin Matrices: New Interactions for Crafting Tabular Visualizations. Charles Perin, Pierre Dragicevic, Jean-Daniel Fekete. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2014), 20(12):2082-2091, 2014.*
- *A Nested Model for Visualization Design and Validation. Tamara Munzner. IEEE TVCG (Proc. InfoVis 2009) 15(6):921-928, 2009.*
- *Criteria for Rigor in Visualization Design Study, Miriah Meyer and Jason Dykes. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2020) 26(1): 87-97, 2020.*
- *Design Study Methodology: Reflections from the Trenches and the Stacks. Michael Sedlmair, Miriah Meyer, and Tamara Munzner. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2012), 18(12):2431-2440, 2012.*
- *Rainbows Revisited: Modeling Effective Colormap Design for Graphical Inference. Khairi Reda and Danielle Albers Szafir. IEEE TVCG (Proc. InfoVis 2020)*
- *D3: Data-Driven Documents. Michael Bostock, Vadim Ogievetsky, Jeffrey Heer. IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis 2011), 17(12), 2301-2309, 2011.*
- *Exploring the Design of Patient-Generated Data Visualizations. Fateme Rajabiyazdi, Charles Perin, Lora Oehlberg, Sheelagh Carpendale. GI (Proc. GI 2020), 362-373, 2020.*

Week-by-week Description:

Week 1: What is visualization? Sketching Visualization Ideas
Definitions of visualization, some advantages of visualization (data analysis), misleading visualizations, sketching visualization designs on paper

Week 2: Data, Tasks, Nested Model, Frameworks
Data types, attributes (Quality, Dimensions), retinal properties of data types, visual encoding, task taxonomy, the nested model for visualization design

Week 3: Rules of Thumb, Design Study Methodology
Principals of good design, design studies in the field of visualization research

Week 4: D3, Tableau, Wrangler
Existing tools, libraries, and software to design visualization, D3: Data-Driven Documents library and JavaScript, Tableau software (interactive data visualization software), Trifacta Wrangler (Data Engineering Cloud)

Week 5: Tables, Bertin Matrices Revisited
Arrange tables, tabular visualization, Bertin matrices

Week 6: Color, Grayscale, Rainbows Revisited, Perception – Applications in Biomedical Engineering
Colour information (schemes, colour-blind friendly schemes, print-friendly colours), sensory discriminability, visual attention, static and moving patterns

Paper*: Evaluation of Artery Visualizations for Heart Disease Diagnosis, Borkin et al.*

Week 7: Student Project Midterm Presentations
Students will have 3 minutes to present their project proposal, progress, plan

Week 8: Interaction/Navigation – Applications in Biomedical Engineering
Interactively changing the data through a visual mechanism (query), dynamic queries, disadvantages of dynamic queries, Fitt's law, brushing and linking, zooming, overview and detail

Paper*: PhenoLines: Phenotype Comparison Visualizations for Disease Subtyping via Topic Models, Glueck et al.*

Week 9: High Dimensionality, Graph, Network – Applications in Biomedical Engineering
Techniques to display multidimensional data (ex. parallel coordinate, Glyph, dimension reduction, small multiples), techniques to display network and graph data visualization (ex. focus and context, reduce items and attributes)

Paper*: ABySS-Explorer: Visualizing Genome Sequence Assemblies, Nielsen et al.*
Paper*: Interactive Visualization of Genealogical Graphs, McGuffin et al.*

Week-by-week Description:

Week 10: Domain-Specific Visualization: Biomedical Engineering
Invited speakers to introduce their innovative visualization systems in the field of biomedical engineering

Week 11: Domain-Specific Visualization: Time-series, Text Data Visualization – Applications in Biomedical Engineering
Techniques to display temporal and periodic data, techniques to display text data

Paper*: Doccurate: A Curation-Based Approach for Clinical Text Visualization, Sultanum et al.*

Week 12-13: Review and Student Project Final Presentations
Discussion of the lesson learnt, fast forward video previews of the projects (30 seconds videos) to showcase the final project outcomes to the whole class

Rationale:

Upon completion of this course, students will be able to:

- *Design interactive information visualization to support human activities using a people-centred process.*
- *Describe visualization design guidelines and explain the guidelines in terms of human perception and cognition.*
- *Critically evaluate information visualization systems, techniques, tools, and research papers.*
- *Describe major sub-disciplines within the field of visualization.*

Currently, there are no other courses offered on Data Visualization in the Department of Systems and Computer Engineering or other units within the OCIBME; thus, there is no direct overlap between this course and any other offered course.

COMP5209/CSI 5135: "Visual Analytics" covers principles, techniques and application of visualization for visual data analysis. Visual analytics is focused on using visualization to analyze data (big data); however, information visualization is focused on how to design, develop, and evaluate data visualization. Thus, the topics in the COMP5209 course do not overlap with the proposed course (BIOM 5403).

ITEC5207: "Data Interaction Techniques" touches on fundamental concepts in data visualization; however, the course is focused on data interaction, data analysis, and visualization design rules. In the BIOM 5403, we will additionally discuss visualization use cases, including the medical and biomedical domains.

OCIBME students are not guaranteed access to the School of IT courses; thus, this course will be offered at SCE. Since there may be some overlap in the introductory parts, this course would preclude credit for ITEC5207.