Syllabus: course M17 CLNV 5303.01, spring 2018
Introduction to Biomedical Informatics II: Methods
Mondays 4 – 7 pm

Course Masters
Preferred method of contact: email
Office Hours: By appointment

Philip R.O. Payne, PhD, FACMI
314-454-5272
prpayne@wustl.edu

Albert M. Lai
314-273-1391
amlai@wustl.edu

Guest Lecturers:

Tiffani Bright, PhD
tjbright@wustl.edu

Randi Foraker, MA, PhD
Randi.foraker@wustl.edu

Tara Payne, MA
Tara.payne@wustl.edu
Philip R.O. Payne, PhD, FACMI
Director, Institute for Informatics
Robert J. Terry Professor
Professor of Computer Science and Engineering

Dr. Payne is the founding director of the Institute for Informatics (I²) at Washington University in St. Louis, where he also serves as the Robert J. Terry Professor and Professor of Computer Science and Engineering. Previously, Dr. Payne was Professor and Chair of the Department of Biomedical Informatics at The Ohio State University.

Background:
Dr. Payne is an internationally recognized leader in the field of clinical research informatics (CRI) and translational bioinformatics (TBI). His research portfolio is actively supported by a combination of NCATS, NLM, and NCI grants and contracts, as well as a variety of awards from both nonprofit and philanthropic organizations.

Dr. Payne received his Ph.D. with distinction in Biomedical Informatics from Columbia University, where his research focused on the use of knowledge engineering and human-computer interaction design principles in order to improve the efficiency of multi-site clinical and translational research programs. Prior to pursuing his graduate training, Dr. Payne served in a number of technical and leadership roles at both the UCSD Shiley Eye Center and UCSD Moores Cancer Center.

Dr. Payne’s leadership in the clinical research informatics community has been recognized through his appointment to numerous national steering, scientific, editorial and advisory committees, including efforts associated with the American Medical Informatics Association (AMIA), AcademyHealth, the Association for Computing Machinery (ACM), the National Cancer Institute (NCI), the National Library of Medicine (NLM) and the CTSA consortium, as well as his engagement as a consultant to academic health centers throughout the United States and the Institute of Medicine.
Research Interests:
- Knowledge-based approaches to the discovery and analysis of bio-molecular and clinical phenotypes and the ensuing identification of precision diagnostic and therapeutic strategies in cancer
- Interventional approaches to the use of electronic health records in order to address modifiable risk factors for disease and enable patient-centered decision making
- The study of human factors and workflow issues surrounding the optimal use of healthcare information technology

Albert M. Lai, PhD
Chief Research Information Officer
Associate Professor, Department of Medicine, Division of General Medical Sciences

Dr. Lai is focused on applying computer science and informatics techniques to solve problems in the clinical domain.

Background:
Albert M. Lai, PhD, is the Chief Research Information Officer for the Institute for Informatics and an Associate Professor of General Medical Sciences at Washington University School of Medicine in St. Louis. Dr. Lai specializes in the development of research informatics infrastructure and is well recognized in the fields of clinical research informatics and consumer health informatics. His recent research has focused on the application of natural language processing, temporal reasoning and information fusion to generate a longitudinal computable phenotype to support clinical trial prescreening. His research portfolio has been supported by a combination of NCATS, NLM, NCI, AHRQ and PCORI grants and contracts. Previously, Dr. Lai served as the Associate Chief Research Information Officer at The Ohio State University Wexner Medical Center.

Dr. Lai received his PhD with distinction in biomedical informatics from Columbia University, where his research focused on designing and evaluating the effectiveness of training older adults to use a home telemedicine system at a distance. Dr. Lai also received his undergraduate and graduate degrees in computer science from Columbia University. While working on his Master’s degree, he conducted research on performance analysis of thin-client computing over wide area networks.

Dr. Lai’s leadership in the clinical research and consumer health informatics communities has been recognized through his election to working group leadership positions within the American Medical Informatics Association as well as his participation in federal study sections and advisory committees.

Research Interests:
- Clinical research informatics
- Clinical informatics
- Consumer health informatics
- Telemedicine
- Usability
- Natural language processing
- Mobile health
About This Course


Journal publications and/or equivalent readings will be posted in Blackboard at least 1-week prior to the lectures for which they are relevant.

Course Description:

This course introduces students to the methods needed in order to apply the foundational theories covered in Biomedical Informatics I. The course will cover a broad spectrum of such methods including both computational and quantitative science techniques that can be employed in the design, conduct, and analysis of basic science, clinical, and translational research programs. This course is intended to enable individuals to critically select such methods and evaluate their results as part of both the design of new project as well as the review of results available in the public domain (e.g., literature, public data sets, etc.).

Core concepts to be reviewed during this course include:

- Foundations of biomedical computation
- Data modelling, codification and integration
- Domain-specific knowledge representation
- Decision science and computational reasoning (including heuristics, logical operation and probability theory)
- Quantitative data analysis
- Data presentation and visualization
- Systems evaluation
- Critical thinking skills surrounding the ability to ask and answer questions about complex and heterogeneous biomedical data

Prerequisite: M17-5302 or instructor permission

Course Objectives:

The objectives of this course are to provide trainees with a solid understanding of core biomedical and quantitative science principles used in the design and execution of basic, clinical, and translational research programs. By the end of the class, trainees should be able to:

- Utilize common computational tools and methods for data capture, storage, and manipulation;
- design, implement, and employ data structures and database management systems suitable for use in the biomedical domain;
- understand and leverage common knowledge representation schemas and constructs in order to annotate or reason upon biomedical data;
- apply in silico hypothesis generation methods to identify and quantify high-order patterns or motifs of interest in complex biomedical data sets;
- apply Bayesian methods to explore probabilistic interconnections between data elements comprising large-scale and complex biomedical data sets;
- use state-of-the-art methods to summarize data and contextualize it relative to targeted samples/populations and observed probabilistic relationships between data of interest; and
- ask and answer questions relative to various types of quantitative relationships between and among data elements making up larger experimental data sets.
Daily Work/Homework:
Each week, students will be expected to turn in a write-up based off of the previous lecture and in-class discussions. Write-ups should incorporate theories and principles learned in the readings, lecture materials, and addressed in the in-class discussions.

Major Assignment Descriptions:
The course will culminate with a written and technical project, as well as an in-person presentation and final exam. This written and technical project will include:

1) The selection of driving biological or clinical problem that can be addressed using biomedical informatics theories and methods as presented in this course;
2) The selection of one or more corresponding public data set(s) that can be used to inform a partial or complete solution to the selected driving problem;
3) The development of a protocol document (2-3 pages) describing how a data analysis script and/or application that utilizes appropriate biomedical informatics theories, applied to the selected public data set(s), can serve to generate one or more data or analytics products that inform a partial or complete solution to your driving biological or clinical problem and how such an outcome will be evaluated;
4) The implementation and evaluation of the data analysis script and/or application as is formalized in the aforementioned protocol document;
5) The preparation of a report (3-5 pages) concerning the outcomes of the preceding implementation and evaluation process, formulated as a case study; and
6) The preparation of a 15m presentation that summarizes the lessons learned from items 1-5 above.

Technology Requirements:
All the class materials, as well as the assignments and communications will be done through Blackboard (found at https://bb.wustl.edu).

Students will be expected to bring a laptop to class for the experiential learning components of the course. Problems with your computer or other technology issues are not a reason for delays in meeting expectations and missed deadlines for the course. If you have a problem, get help in solving it immediately. At a minimum, you will need the following software/hardware to participate in this course:

1. A laptop computer with a current operating system (e.g. Windows, Mac OS, Linux)
   a. You will need appropriate privileges on this computer to be able to install additional, open-source software packages as assigned by the instructors throughout the class
2. An up-to-date internet browser (preferably Google’s Chrome or Apple’s Safari)
3. An active WUSTL key (such that you can log into the wireless network and/or other university computing resources)
4. The ability to access and use the Blackboard Learning Management System (https://bb.wustl.edu)
5. Access to a high-speed internet connection

In addition, all students will be expected to obtain a username and password for the National Library of Medicine’s (NLM) terminology services. This requires the completion of an account request and user license at the following web-site: https://uts.nlm.nih.gov/home.html
Time Requirements

For face-to-face courses in the CRTC program, it is expected that you will be in class 1 hour per week for each credit of the course a week plus travel time (i.e. this is a 3-credit course so that is 3 hours a week). In addition, it is assumed you will be doing homework and reading assignments that take at least double that time. You should anticipate your time commitment for this course to be at least 9 hours a week.

Course Schedule (subject to modification)

<table>
<thead>
<tr>
<th>Dates</th>
<th>Topic</th>
<th>Instructor(s)</th>
<th>Chapters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/22</td>
<td>Course Logistics and Overview &amp; Computation in Biomedicine</td>
<td>Lai</td>
<td>Chapter 1</td>
</tr>
<tr>
<td>1/29</td>
<td>Database design, integration and metadata management</td>
<td>T. Payne</td>
<td>Chapter 2, Appendix C</td>
</tr>
<tr>
<td>2/5</td>
<td>Informatics Methods Supporting Personalized Medicine</td>
<td>P. Payne</td>
<td>Chapter 11</td>
</tr>
<tr>
<td>2/12</td>
<td>Knowledge Representation and Knowledge-Based Systems Design</td>
<td>P. Payne</td>
<td>Chapter 3, Chapter 7</td>
</tr>
<tr>
<td>2/19</td>
<td>Linking Data Across Scales</td>
<td>P. Payne</td>
<td>Chapter 12</td>
</tr>
<tr>
<td>2/26</td>
<td>Reasoning (Heuristics, Probability, and Bayes Theorem)</td>
<td>Foraker</td>
<td>Chapter 8, supplemental readings</td>
</tr>
<tr>
<td>3/5</td>
<td>Hypothesis Generation and Classifiers</td>
<td>Swamidass</td>
<td>Chapter 4, Chapter 9</td>
</tr>
<tr>
<td>3/12</td>
<td>NO CLASS, SPRING BREAK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/19</td>
<td>Data Summarization</td>
<td>Foraker</td>
<td>Assigned Readings</td>
</tr>
<tr>
<td>3/26</td>
<td>Asking and Answering Questions About Data</td>
<td>Foraker</td>
<td>Assigned Readings</td>
</tr>
<tr>
<td>4/2</td>
<td>Data Visualization</td>
<td>Yen</td>
<td>Assigned Readings</td>
</tr>
<tr>
<td>4/9</td>
<td>Systems Evaluation</td>
<td>Payne/Yen</td>
<td>Assigned Readings</td>
</tr>
<tr>
<td>4/16</td>
<td>Software Engineering and Technology Deployment</td>
<td>Lai / Payne</td>
<td>Chapter 10</td>
</tr>
<tr>
<td>4/23</td>
<td>Review/Discussion</td>
<td>Payne</td>
<td>Chapter 13</td>
</tr>
<tr>
<td>4/30</td>
<td>Final Projects Due and In-Person Presentations</td>
<td>Payne</td>
<td></td>
</tr>
<tr>
<td>5/7</td>
<td>Final Exam</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Assessment/Grading

Grade Composition:

Summary of Course Assignment Point Values:

Weekly Lecture Write-ups (grades as Pass/Fail)) 15%
One page topic proposal
Final Project (25% paper/15% presentation) 40%
Final Exam 30%
Attendance, Participation, and Professionalism 15%

100

Grading Scale:

<table>
<thead>
<tr>
<th>Grades/sub-grades</th>
<th>Course Points</th>
<th>4-point scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+ (98% to 100%)</td>
<td>98-100</td>
<td>4.00</td>
</tr>
<tr>
<td>A (93% to 97%)</td>
<td>93-97</td>
<td>4.00</td>
</tr>
<tr>
<td>A- (90% to 92%)</td>
<td>90-92</td>
<td>3.7</td>
</tr>
<tr>
<td>B+ (88% to 89%)</td>
<td>87-89</td>
<td>3.3</td>
</tr>
<tr>
<td>B (83% to 87%) – minimum for Core courses</td>
<td>83-86</td>
<td>3.00</td>
</tr>
<tr>
<td>B- (80% to 82%)</td>
<td>80-82</td>
<td>2.7</td>
</tr>
<tr>
<td>C+ (77% to 79%)</td>
<td>77-79</td>
<td>2.3</td>
</tr>
<tr>
<td>C (73% to 77%) – minimum for Electives</td>
<td>73-76</td>
<td>2.00</td>
</tr>
<tr>
<td>C- (70% to 72%)</td>
<td>70-72</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Penalties for Late Work:
Late work will not be accepted, unless in the event of unforeseen circumstances. If these situations arise, students must receive written approval from the coursemaster. If you must miss a class, please email a copy of the weekly assignment to Andrea Krusel prior to the start of the class when it is due.

Attendance Requirement:
In-class participation is an important part of the coursework taken as part of the MSCI or AHBR programs and the clinical research training programs within the CRTC. Students are expected to physically attend at least 75% of class sessions for each course they take. Watching the videotaped class presentations, if available, is helpful to keep up with missed sessions, but is not a substitute for class attendance. Students whose professional duties or personal circumstances prevent them from meeting this program attendance requirement must receive prior written approval of the coursemaster, and agree on an alternate plan to achieve course objectives and earn academic credit.

Blackboard Support
If you have any technical problems accessing Blackboard please e-mail crtc@email.wustl.edu. Note, this mailbox is not monitored in the evening or on weekends. If you need immediate help after hours please put a service request into https://wusm.service-now.com.

Course Policies

Participation:
• It is vitally important that our classroom environment promote the respectful exchange of ideas. This entails being sensitive to the views and beliefs expressed during discussions whether in class or online.

• Your success in this course will heavily depend on your ability to communicate, engage and participate in all course activities. Successful completion of this course requires that a student keep up with all assignments and prep work for the lab components.

If you are unable to participate in the scheduled class activity or discussions you must notify the coursemaster within the week of that class module or discussion. **An unexcused failure to engage or participate with the class will be counted as an absence; unexcused absences may result in failure.** The coursemaster reserves the right to make judgment to accept and/or make–up assignments missed because of failed participation in the course activities.

**Drop Dates:**
If the occasion should arise that you want or need to drop this class, please see me first. You can drop for any reason during the course of the semester, however you my only receive a partial or no tuition reimbursement depending upon how far into the semester you drop the course. See the Academic Calendar for your program for specific dates and reimbursement policies. Note, late withdrawals will also appear on your transcript as a withdrawal.

**CRTC Academic Policy Guidelines:**
Guidelines regarding CRTC course registration and enrollment, grades, tuition obligation, and academic leave are consolidated in the [CRTC Academic Policy Guidelines](#). Please take a moment to review this document.

**CRTC Guidelines for Academic and Non-Academic Transgressions:**
By registering for this course, you have agreed to the terms of the [CRTC Guidelines for Academic and Non-Academic Transgressions](#). If you have not already reviewed this policy, please be sure to before beginning any CRTC related coursework.

**Academic Integrity/Plagiarism:**
• Academic dishonesty is a serious offense that may lead to probation, suspension, or dismissal from the University. One form of academic dishonesty is plagiarism – the use of an author’s ideas, statements, or approaches without crediting the source. Academic dishonesty also includes such acts as cheating by copying information from another student. **Plagiarism and cheating are not acceptable.**

• Academic dishonesty will be reported to the Office of the Registrar for possible action. The coursemaster will make an academic judgment about the student’s grade on that work and in that course. The CRTC process regarding academic dishonesty is described in the [CRTC Guidelines for Academic and Non-Academic Transgressions](#)

**Writing Assistance:**
For additional help on your writing, consult the expert staff of [The Writing Center](#) in Olin Library (first floor). It can be enormously helpful to ask someone outside a course to read your essays and to provide feedback on strength of argument, clarity, organization, etc.

**Mental Health Resources:**
Mental Health Services’ professional staff members work with students to resolve personal and interpersonal difficulties, many of which can affect the academic experience. These include conflicts with or worry about friends or family, concerns about eating or drinking patterns, and feelings of anxiety and depression. See: [http://shs.wustl.edu/MentalHealth](http://shs.wustl.edu/MentalHealth)

**Reporting Policies:**
Please also review the CRTC website for policies regarding sexual assault reporting and reporting concerns about bias, prejudice or discrimination.