Example Syllabus for Independent Study on Explainable AI
with Prof. Iris Howley (last updated February 2020)

This project explores the relationship between instructor and student understanding of the artificial intelligence (AI) algorithms that underlay their educational technology, and the impact of that algorithmic understanding on decision-making for learning. This independent study will involve the development of "explainables" or brief, engaging interactive tutoring systems to provide algorithmic understanding to classroom stakeholders. The explainables developed by this independent study will be publicly accessible and usable by external projects, increasing algorithmic understanding for the initially intended stakeholders, but also for the general public. This work combines approaches from the learning sciences, human-computer interaction, ethics, and machine learning.

Continued work on this topic will involve studies with people to investigate how algorithmic understanding impacts system trust and decision-making for learning, as well as the necessary knowledge for “understanding an algorithm.” The overall goal of this line of work is to yield a framework for designers of algorithmically enhanced learning environments to determine what level of algorithmic understanding is necessary to achieve the goals of informed decision-making by users of their systems. The main contributions of this work include a methodologically rigorous investigation of the knowledge components of algorithmic understanding for learning contexts that can be applied to model interpretability discussions in the wider machine learning community. The contributions of this particular independent study include adding to ongoing discussions about ethical algorithmic transparency in the larger machine learning community, providing an actionable framework for developing a more AI-informed student and teacher body as well as building lightweight explainables for appending to external algorithmically enhanced learning environments.

The final deliverable for this project will be a paper and a final project that is a web-accessible explainable for Bayesian Knowledge Tracing (an AI algorithm). The paper will describe this semester’s process, including: a summary of the existing research, analysis of existing projects, a high-level view of the steps of this project and why decisions were made, relevant images and evidence, and a summary of the technical implementation. Each week, the readings and deliverables will be approximately 12 hours of work, and I will meet with Iris for ~1 hour to discuss the weekly topic.

Week 0. Preparation

Guiding Questions: What are the main goals of this work? Are they good goals? How will we achieve them?

Reading:
1) Iris’ NSF CRII Grant Proposal

Deliverables:
1) Join the HCI Research Slack group
2) Complete the RCR and HSR training on CITI (email Iris the certificates)

Week 1. Learning & Teaching with BKT

Guiding Questions: What do Bayesian Knowledge Tracing Systems do? What additional benefits do they provide for instructors and students? What are some of the practical concerns of using a BKT system?

Reading:

Deliverables:
1) Pick-up a web accessible programming language:
   a. Begin a Javascript Tutorial (maybe this React tutorial on Scrimba)
   b. …or begin a D3 Tutorial (Emma Saunders’ D3.js Essential Training for Data Scientists)
2) Begin this semester’s paper by starting a section titled “Prior Work.” Summarize this week’s readings.

Week 2. Explainability Goldilocks

Guiding Questions: Is transparency good? What is an effective explanation? How has past work provided the right level of AI explanation to improve users’ lives?

Reading:

Deliverables:
1) Continue learning a web accessible programming language. Be prepared to demo what you’ve built so far:
   a. Continue the Javascript Tutorial (maybe this React tutorial on Scrimba)
   b. …or continue the D3 Tutorial (Emma Saunders’ D3.js Essential Training for Data Scientists)
2) Continue adding to the literature review document. Summarize this week’s readings.
Week 3. BKT Deep Dive

**Guiding Questions:** What is Bayesian Knowledge Tracing? How does the algorithm work?

**Reading:**


3) Choose (at least) one:

**Deliverables:**

1) Continue learning a web accessible programming language. Be prepared to demo what you’ve built so far:
   a. Continue the Javascript Tutorial (maybe [this React tutorial on Scrimba](https://www.scrimba.com/courses/react/
   b. …or continue the D3 Tutorial (Emma Saunders’ [D3.js Essential Training for Data Scientists](https://www.youtube.com/watch?v=L1FyDy59KXg))

2) Continue adding to the literature review document. Summarize this week’s readings.

---

Week 4. Explainables Exploration

**Guiding Questions:** What makes a good explainable? What features of an explainable make you want to use it more? What features help you learn best? What features are the most fun?

**Reading:**

1) Ko. "*How to be Creative,*" *Design Methods.*

2) Interact with existing Explainables and note what you like and don’t like. Which one did you learn the most from?:
   a. Carter &., 2017. "*Using Artificial Intelligence to Augment Human Intelligence,*" distill.pub
   b. Yee & Chu, 2015. "*A visual introduction to machine learning,*" *R2D3 (Personal Website).* And [Part II](https://www.r2d3.org/
   c. Carter et al, 2016. "*Four Experiments in Handwriting with a Neural Network,*" distill.pub. See the [github repo](https://github.com/colson32/) available.
   e. Bostock, 2015. "*Visualizing Algorithms,*" *Personal Website.* (Uses D3). The #Related Work section at the bottom has more model explainables
   g. Strobelt et al, 2018. "*Art - Inspired Data Experiments on Neural Network Model Decay,*" *Forma Fluens Project Website* and the [Forma Fluens Main Page](https://www.formafluens.com/).
h. Knight et al, 2018. "Evidence Based Teaching Guide: Peer Instruction," CBE Life Science Education. (This is an Explainable, but not for AI. A decision-graph approach).

i. The building blocks of interpretability.

j. Explained Visually.

k. "Explorables" that explain many other concepts: explorable.es.

**Deliverables:**

1) Come to the meeting with 20 one-sentence ideas for your own explainable

2) Continue learning a web accessible programming language. Be prepared to demo what you’ve built so far:
   a. Continue the Javascript Tutorial (maybe this React tutorial on Scrimba)
   b. …or continue the D3 Tutorial (Emma Saunders’ D3.js Essential Training for Data Scientists)

3) Add a new section to the research paper titled “An Analysis of Existing Explainables.” Compare and contrast some of the most successful features of explainables explored this week. Take an evidence-based approach by including links and screenshots of existing explainables.

---

**Week 5. Sketching Ideas**

**Guiding Questions:** What are the purposes of sketches and storyboards? How much detail should be included in a sketch? How much detail in a storyboard?

**Reading:**


3) A textbook to refer to: Buxton et al, 2007. Sketching user experiences : getting the design right and the right design. This book is available for e-access from the Williams library (and one physical copy, T359 .B89 2007- might be in TCL312)

**Deliverables:**

1) Choose 3 of the ideas from last week to sketch out a slightly deeper design using storyboards or sketches. Who will use this design? When?

2) Continue learning a web accessible programming language. Be prepared to demo what you’ve built so far:
   a. Continue the Javascript Tutorial (maybe this React tutorial on Scrimba)
   b. …or continue the D3 Tutorial (Emma Saunders’ D3.js Essential Training for Data Scientists)

3) Add a section to the research paper titled “Design Methods.” Include images of your sketches, why you chose them, and any references that would be helpful. Also add an “Introduction” section to the research paper, and include a first draft of the goals of your project.
Week 6. Paper Prototyping

Guiding Questions: Why paper prototype? What insights do paper prototypes provide? Why paper prototype before building a full implementation? How do you create a paper prototype?

Reading:


Deliverables:

1) Build a paper prototype (or PowerPoint prototype) of one of your ideas from the previous week.
2) Continue learning a web accessible programming language. Be prepared to demo what you’ve built so far:
   a. Continue the Javascript Tutorial (maybe this React tutorial on Scrimba)
   b. …or continue the D3 Tutorial (Emma Saunders’ D3.js Essential Training for Data Scientists)
3) Similar to the previous week, add an evidence-based summary of your paper prototyping process.

Week 7. Usability Testing

Guiding Questions: What are the steps of running a usability test? What amount of detail in the task is ideal?

Reading:

2) Snyder, Ch. 8 "Introduction to Usability Test Facilitation", pp. 171-195, in Paper Prototyping.

Deliverables:

1) Run a usability test with at least three people, using your paper prototype from the week before. What went well, what didn’t? What do you need to fix? What were your tasks?
2) Continue learning a web accessible programming language. Be prepared to demo what you’ve built so far:
   a. Continue the Javascript Tutorial (maybe this React tutorial on Scrimba)
   b. …or continue the D3 Tutorial (Emma Saunders’ D3.js Essential Training for Data Scientists)
3) Add a section to your research paper titled “Usability Tests” and provide a list of feedback you received from users (direct quotes are really helpful), as well as steps you will take to address those issues.
Week 8-11. Project Implementation

**Reading:**
1) Assorted readings, as needed, relevant to your project

**Deliverables:**
1) Implement your project in a web accessible format.
2) Add a section to your research paper titled “Implementation” and describe the tools used to develop your final deliverable.

Week 12. Project Communication

**Guiding Questions:** Why is communicating our ideas effectively important? What makes a good elevator pitch?

**Reading:**
1) What are the elements of an effective elevator pitch?

**Deliverables:**
1) Bring a 5 second, 30 second, and 2 minute elevator pitch for your research and be prepared to give and improve them.
2) Create documentation for your project that includes how to set-up and run your project on new machines (for new students).
3) Complete the “Introduction” section of your research paper. Also add a “Conclusion” section where you summarize the main contributions of this work.

Week 13. Project Submission

**Deliverables:**
1) Provide a link (or a github repo) to all the files necessary to set-up and run your project. Include the documentation.
2) Share the final research paper with Iris.

Optional Reading Modules:

Week ??, Cognitive Task Analysis

**Guiding Questions:** What are the steps of a Cognitive Task Analysis? What insights does a CTA provide that a semi-structured interview doesn’t?

**Reading:**