

# Scott Cambo

Northwestern University, PhD Student  
cambo@u.northwestern.edu  
www.scottallencambo.com



Human-centered machine learning, machine teaching,  
model explanations, domain adaptation

## Interactively Tailoring Machine Learning for Situated Use

When building systems that use AI or machine learning, it is often difficult or even theoretically impossible (e.g. *No Free Lunch theorem*) to avoid making assumptions about the data in a real world application. When these assumptions fail, users are often left without an understanding or recourse for correction. My work addresses three challenges in designing systems that will help users assert their situated perspective and domain expertise on the models that affect the user's ability to manage their health, their work life, and their participation with modern culture.

### 1. Model Explanations

Modern machine learning algorithms have grown in complexity at the expense of transparency (e.g. Deep Learning). However, frameworks like Local Interpretable Model-agnostic Explanations permit new approaches to communicate the algorithmic process in addition to the training data. My work aims to understand which algorithmic mechanisms, methods for discovering them, and methods for communicating them are most helpful for the tasks users want to accomplish.

### 2. Machine Teaching

The research and practice of machine learning typically focuses on optimization and design of the learning algorithm. Machine Teaching has been proposed as a new area of research to understand how we can design systems that allow non-expert users to “teach” or add knowledge to the system. You can read more about how I plan to contribute to the research area of machine teaching in the last section of this poster.

### 3. Domain Adaptation

In order to make the best use of a large and general training dataset while still getting good accuracy for specific use cases, we can use a set of algorithms for domain adaptation or transfer learning. One of the biggest challenges for domain adaptation in deep learning is that it is difficult to know which part of the model learned concepts that are appropriately general and which are the specific parts that should be adapting with context. My work aims to use model explanations and machine teaching to help users guide the domain adaptation process.

## Personalizing Activity Trackers

*How can seamful design help us communicate model uncertainty in activity tracking interfaces?*

When users engage with tracked information about their health activities, they are not simply using it to track and adjust their health related behaviors. They are also developing an understanding of how the system works and how they can use it to achieve their goals. I am currently studying how we can use **seamful design** so that model certainty is expressed in the interface judiciously.

*How can a mixed initiative design help users adapt general activity tracking models to their specific needs and values?*

**Active Learning** has been proposed as a way of integrating user knowledge into activity tracking systems by eliciting feedback during moments of model uncertainty. My work builds on recent contributions in **mixed initiative design** with intelligent systems to understand how activity tracker interfaces can help users to proactively contribute helpful data in an active learning based activity tracking system.



*“Walking” can seem like an objective concept which should pose little challenge to the machine learning models used in activity tracking systems. However, users often find limitations in seemingly normal settings like hiking over uneven terrain, walking while dragging a bag, walking with the use of a cane, or walking while holding hands.*

## Detecting Inappropriate Behavior Online

*Is what you call “inappropriate”, the same as what I call “inappropriate”?*

Over the last few years, research in CSCW and CHI have highlighted the nuances of how we perceive and label concepts like “inappropriate” content online. My work aims to understand how these communities converge on a practical definition of these concepts and how that may differ from other online communities and those developing intelligent systems for managing inappropriate content.

*How can we identify which aspects of a concept like “inappropriate” are universal and which aspects need to be situated in a specific community and task?*

One of the biggest challenges in **domain adaptation** is understanding what parts of the model should be rigid and which should adapt as we train models for specific contexts. I hope to design new approaches to domain adaptation which appropriately adjust to the needs of each community.



*Online communities face many different types of toxic and inappropriate content that can be highly contextualized in the communities' identity and purpose for forming. Automated content management can use **Machine Teaching** to better utilize explicit input from community members in this situated context.*

An Open Source Curriculum for Human-Centered Machine Learning  
[https://github.com/scottofthescience/hcml\\_course](https://github.com/scottofthescience/hcml_course)

With the help of Professor Bryan Pardo, I am designing an open-source curriculum for teaching Human-Centered Machine Learning. The current version is designed as a graduate seminar to be taught at Northwestern University in Winter 2019 drawing from research in AAAI, NIPS, KDD, IJML, IUI, CHI, CSCW, and more. We encourage our colleagues from all over to suggest papers by submitting pull requests and issues and to fork the curriculum repository for new versions.