Designing AI Applications

Prof. James A. Landay
Computer Science Department
Stanford University

Summer 2018

*Based on *The UX of AI* by Josh Lovejoy and *Human-Centered Machine Learning* by Jess Hollbrook and Josh Lovejoy

Outline

- Principles of human-centered AI
- Examples

Principles

- Address human needs
- Ensure trust by exposing underlying system goals
- Fake it until you make it
- AI apps are socio-technical systems
- Designing for augmentation, not automation

Address Human Needs

Or, Who Needs a Smart Toaster?

- Machine learning will not guarantee it is *the right problem* to solve!
- Even more important to use design thinking process to find a real human need

Ensure Trust by Exposing Underlying System Goals

- If it is unclear how user's interaction fits into system (e.g., how it learns), they will create their own story
- Ensure trust by exposing underlying system goals
- Machine learning is brittle
  - often train many underlying classifiers for every issue you care about. User may not understand this.
  - e.g., in a system to find good photos, need to also train a classifier to get rid of bad ones (e.g., hand in front)

Ensure Trust by Exposing Underlying System Goals

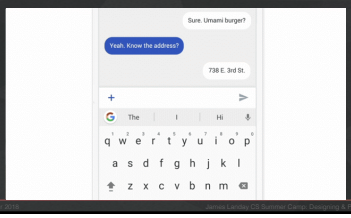
- Understand which failures you can handle

		Prediction: Positive	Prediction: Negative
Reference	Positive	True Positive	False Negative
Reference	Negative	False Positive	True Negative

- Confusion matrix let's you reason about types of error
- Different apps will optimize for different cells of this matrix
  - can you deal with true things being left out? (false negative)
  - is making a false classification worse for you? (false positive)

Fake it Until You Make it

- Getting AI models right takes time, data, & adjustment
- Better to prototype your UX with *real user content*
- Wizard of Oz things that don't yet work



AI Apps are Socio-Technical Systems

- AI systems trained on data collected from people
- Labels used to describe the data is from people
- Did the system do the right thing? People decide

- Human judgement in all of these decisions. Need to include human expertise in any major AI project.

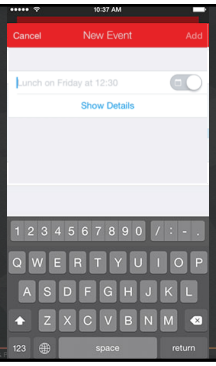
Design for Augmentation, Not Automation

- Machine learning will always make errors
- Key is to help make people more powerful in what they do, rather than try to fully replace them

Example: Surroundview

Example: Fantastical

NLP (text or speech) drives filling out of calendar entries



Example: gmail suggested auto responses