Motivation
- 10-11 Mil. Visually Impaired People in U.S.
- 45+ Million People Without Sight Worldwide
- At least 50% of Blind Americans Use Internet
- The problem with screen readers:
  - narrate web pages sequentially resulting in severe information overload which makes web browsing • Time-Consuming • Strenuous
  - This is more prominent when filling out web forms and conducting web transactions

Label Association Problem
- Use the 'label' attribute
  - 89% form fields don’t have any 'label' attribute
- Take the preceding text in the HTML DOM tree
  - preceding text may not be the target label

Our Approach
- Finite Mixture Models
  - Label of a form element can be generated from any of the candidate labels
  \[ p(d_i|\theta) = \sum_{k=1}^{M} p(k|\theta)p(d_i|k) \]

Expectation-Maximization
- E step
  - Compute expected labels:
  \[ p(k|d_i) = \frac{p(d_i|k)p(k)}{p(d_i)} \]

- M step
  - Update model parameters:
  \[ p(d_i|k)p(k) \rightarrow p(d_i|k)p(k) \]

Labeling Transaction-centric objects
- Non-visual web transactions are difficult
  - Need to locate relevant buttons and links
  - Find context of transaction-centric web elements
  - Follow missing label identification algorithm

Experimental Evaluation
- Effectiveness of Label Association
- Comparison with LabelEX [VLDB’08]

Interface and User Study
- User Interface: HearSay
- A non-visual web browser: provides basic screen-reading functionalities
- Tasks: Searching for jobs, renting cars, reserving flights
- Avg. time to complete each task

This material is based upon work supported by the National Science Foundation - Awards: IIS-0534419, CNS-0751083 and IIS-0808678