Data quality is very important
The Data Entry problem

- Data entry is the first opportunity to address data quality
- Form design is often an ad hoc process
- New entry modalities {web, mobile, touchscreen} require customization
Automatic Form Generation

- Given set of form fields and sample values
- Want “black box” form generator to give
  - Question ordering
  - Widget choice
  - Widget “decorations”
Collaborative Design & Entry

- Allow users to comment on each form field
  - For design feedback: “This field should come later”
  - For date entry comment: “There is a smear, I think this is the correct value”
- Aggregate previous comments per field
- Show based on time or relevance
Visualizing Collaborative Data Entry Forms

- Algorithm for dynamic layout and pagination
  - Consider questions + widgets + “decorations” + comments
  - On web, mobile, touch-screen interfaces
Visualizing Collaborative Data Entry Forms, cont.

● How to represent “risk” of a current value
  ● Example:
    ● What is your gender? M / F
    ● Are you pregnant? Yes / No

● How to represent prior values?
  ● Example 1
    ● M (50%) - F (50%)
  ● Example 2
    ● Given Gender = M
    ● Pregnant (1%), Not Pregnant (99%)
Widget decorations!

- Type-ahead
- Audio signal
- Screen flash
- Pop-up warning
- On/off tabbing between fields
- Label color, size, position
- Alignment, i.e. differently-justify at-risk data
- On/off edit-ability
- Background color
- Entered data color, font, size
- Default values
- Show prior entries
- Prior entry statistics
- Sparkline visualizations
Milestones

- Generate form from XML specification
- Design “decorations” for form widgets
- Build decorations
- Design layout algorithm
- Implement algorithm
- Visualize Tanzania AIDS treatment forms
- Simple evaluation
- Write-up
Related Work

http://vis.berkeley.edu/courses/cs294-10-fa08/wiki/index.php/FP-KuangChen