

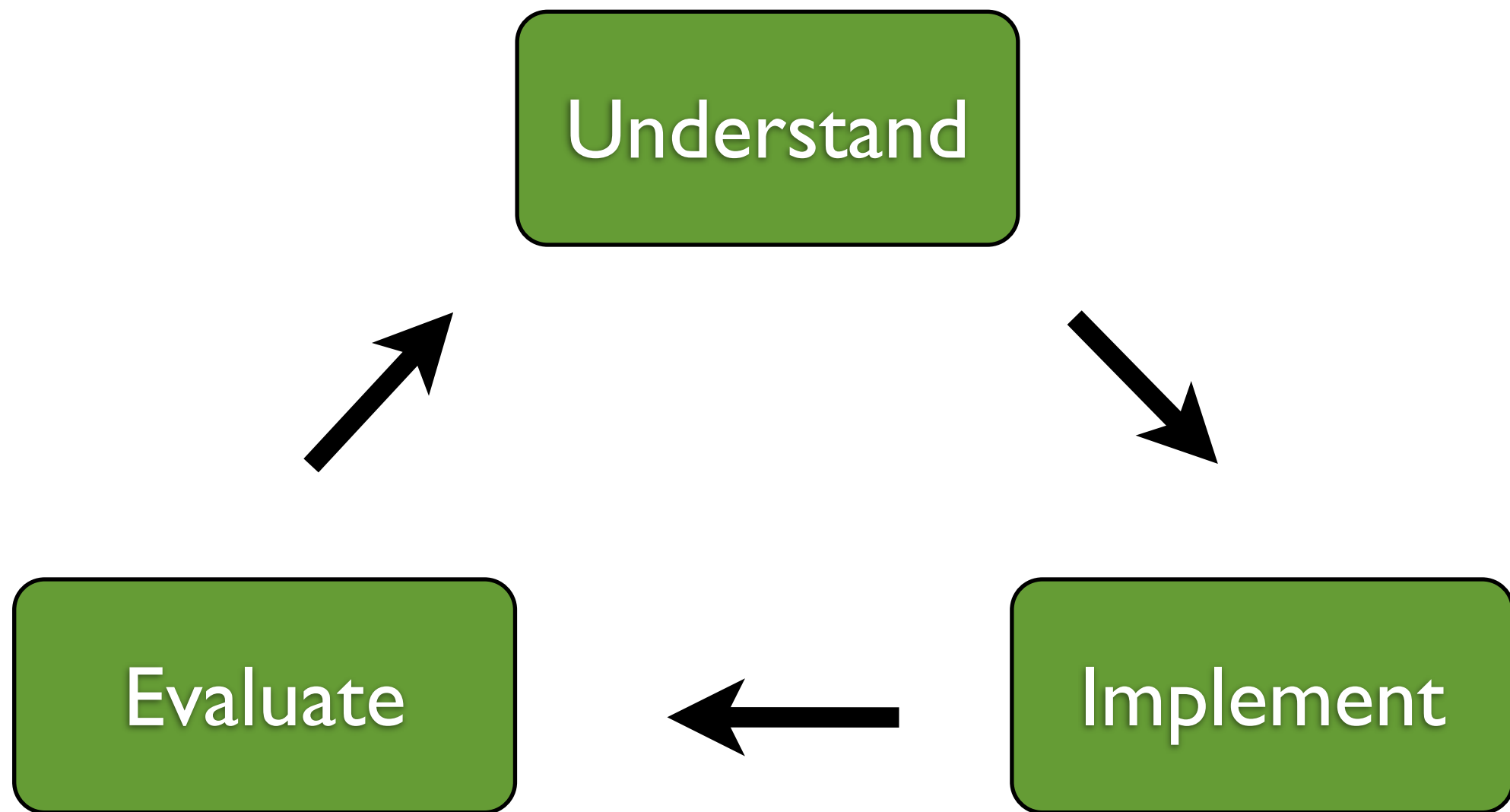
Development as Design

tapan parikh • parikh@berkeley.edu

Development and Learning

- A toolkit of useful methods
 - RCTs, econometrics, ethnographic methods, surveys, large-scale data analysis, etc.
- But still not an integrated process
- *Design* provides a framework for iterative, data-driven problem-solving

Development as Design



Information technology not only affords new solutions - but new ways to understand problems, and to monitor and evaluate solutions (e.g. institutions, technologies, policies)

- **Avaaj Otalo** - Question and Answer for Small Farmers in India
- **Usher** - Dynamic Forms for Data Capture

Agriculture and Development

- The majority of the poor make their living from agriculture
- Improving productivity and profitability of agriculture is the pathway for development of poorest countries (WDR 2008)
- Doing this, while minimizing ecological risk, requires precise information and planning



Agricultural Extension

- Farmers have many questions
 - New techniques, technologies
 - Treating specific pests?
 - Amount, type of inputs to use?
- Extension programs are costly, and don't reach most farmers
- Only accessible resource is local input dealer



Mobiles and Voice

- Mobile phones are reducing the physical limitations of access
- However, information must still be accessible, trusted and relevant
- Local sources are more relevant, credible
- Voice content can be used and created, by small farmers with low-cost handsets

Speaking the language



தர்ந்தெடுக்கப்பட்ட ள்ளார்

தேதி : 12/10/02

1		விஜயா	50	50	 4 பிரியா	 50		
2		பாரதிதாசன்	50	50				
3		கார்த்திக்	50	0				
4		பிரியா	50	<input type="checkbox"/>				
5		மீனா	50		12.10.02	50	✓	
6		லஸ்வி	50		05.10.02	50	✓	
7		சித்ரா	50		22.09.02	50	✓	
8		சத்யானந்தன்	50		15.09.02	0	✗	
9		சகுந்தலா	50		03.09.02	50	✓	
10		வீவகானந்தன்	50		26.08.02	50	✓	
11		கார்த்திக்	50					

பூ ஆரம்ப இருப்பு 12.10.02

குழுவின் மொத்த சேமிப் 1080

பூ Group Investments

பூ கையிருப்பு 1600

பூ Bank balance 6850



Design for Low text literacy

- Leverage existing representations
- Use icons that are familiar and realistic
- Provide guidance throughout the task
- Numbers are more accessible than text
- Local language audio is very important

Orality (Ong, 1982)

- Oral communities have their own ways of representing and managing information
- Aggregative - tolerant of repetition, redundancy and inconsistency
- Situational - tied to specific situations and people; not abstract concepts
- Dialectic - reinforced by dialogue

Avaaj Otaalo









Avaaj Otalo

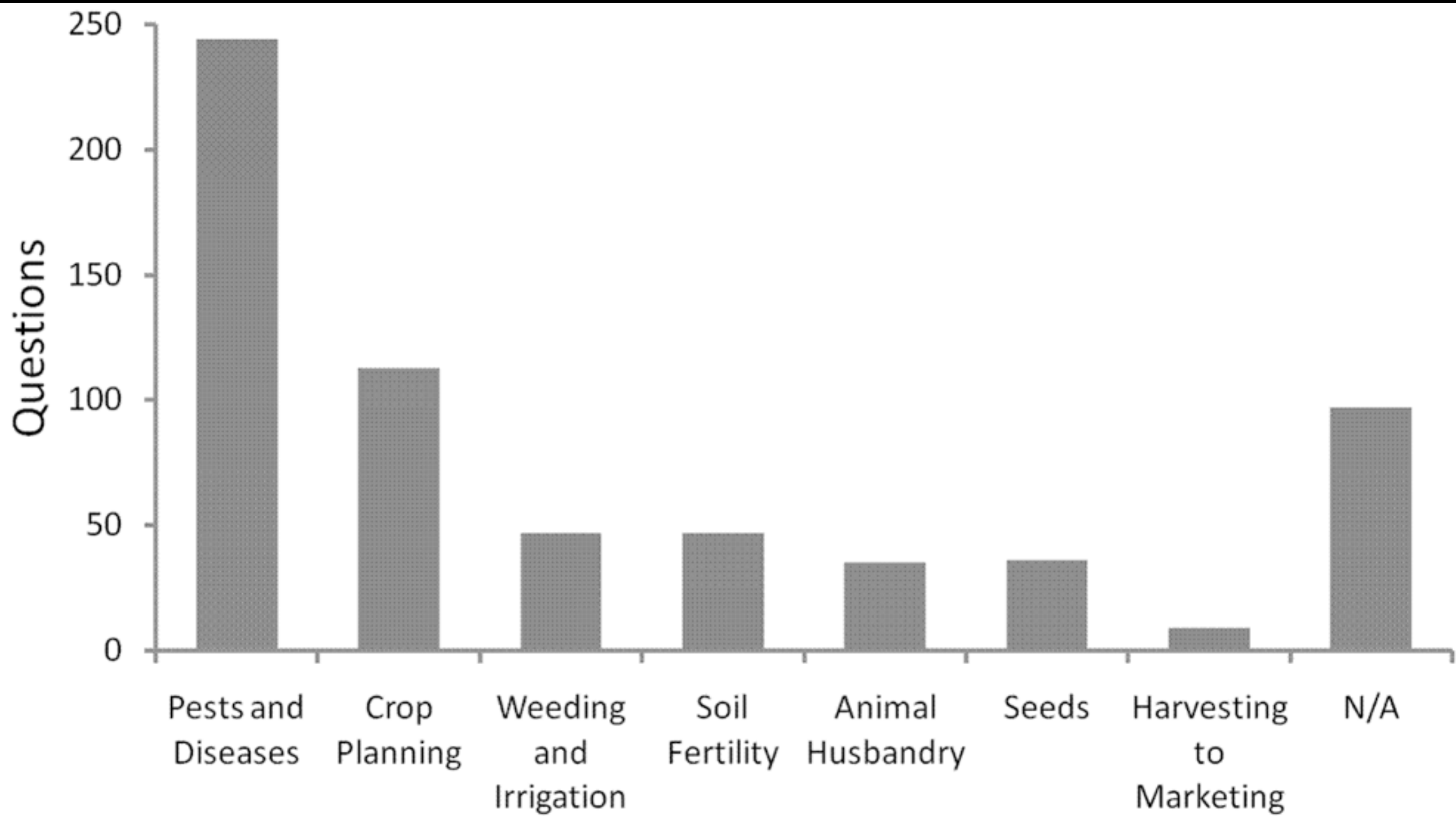
- Usenet using Voice
- Farmers and experts call an IVR-based voice system to:
 - record questions
 - provide answers
 - review previous questions and answers
- Popular questions broadcast on radio

Neil Patel



Current Status

- Pilot with 50 users from Dec 2008 to July 2009 (currently statewide)
- Averaged 1000 calls per month
- One farmer self-reported a \$6K increase in income due to AO!
- Questions covered a variety of agricultural and non-agricultural topics



Farmers Need...

- Knowledge about inputs
- Solutions for pests and diseases
- Access to markets, buyers
- Weather forecasts
- Financial services
- A chance to sing!

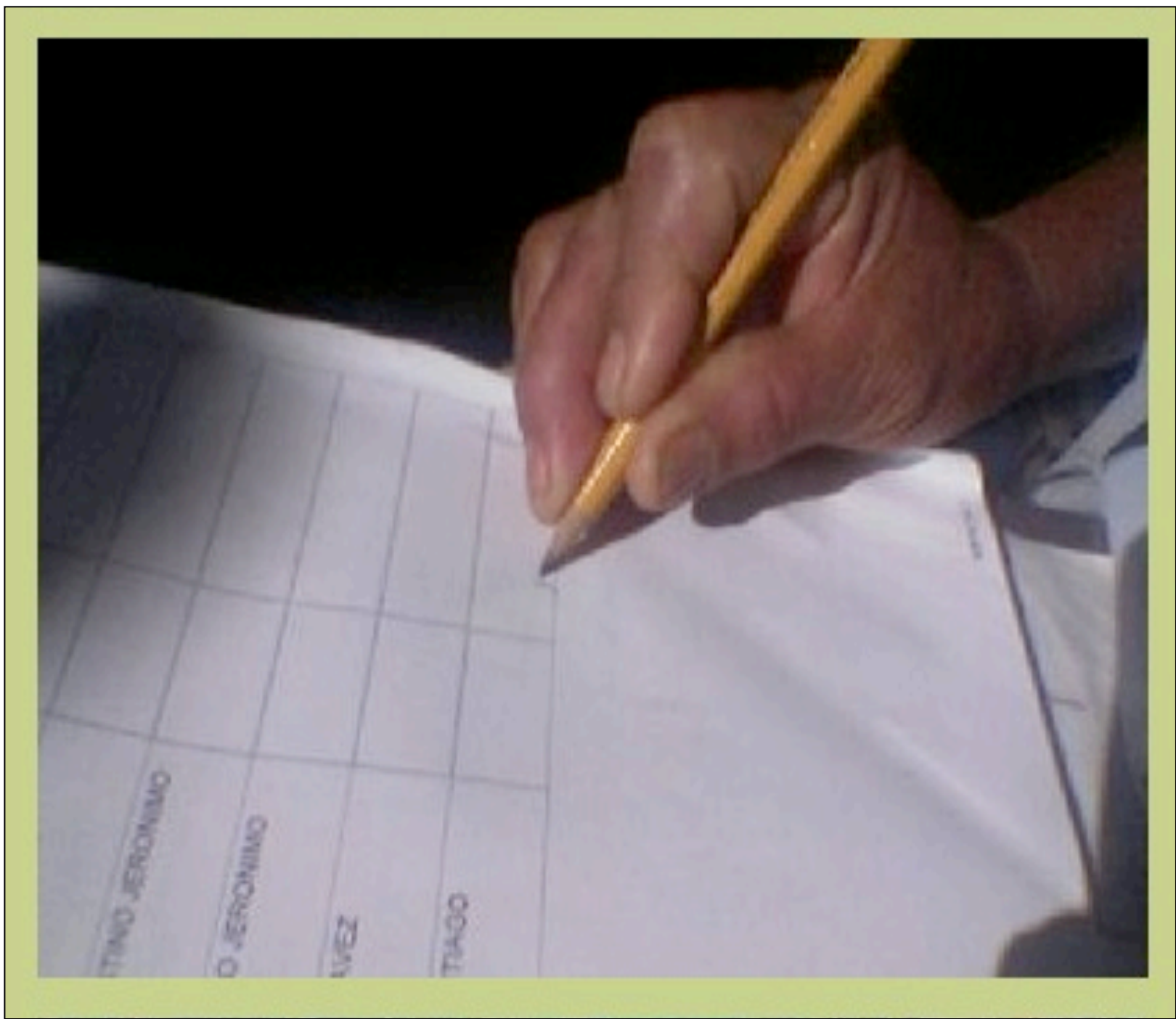
Asking the right question

What problem are you facing?

Asking the question right

Voice is an accessible medium

Collecting data



Data Collection

- Much of the data for evaluation is collected by enumerators, using paper forms
- Surveys can often be very long and tedious (> 300 questions, 3 hours)
- Small orgs lack expertise in form design
- Paper transport, transcription delays

Mobile Phones

- Mobile phones can improve the efficiency and accuracy of data entry
 - Capture once and remotely transmitted
 - Built-in validation
- CAM was the first general purpose tool for data collection using mobile phones
 - JavaRosa, ODK (Android)





Courtesy of Open Data Kit - <http://www.opendatakit.org>

Many Different Forms

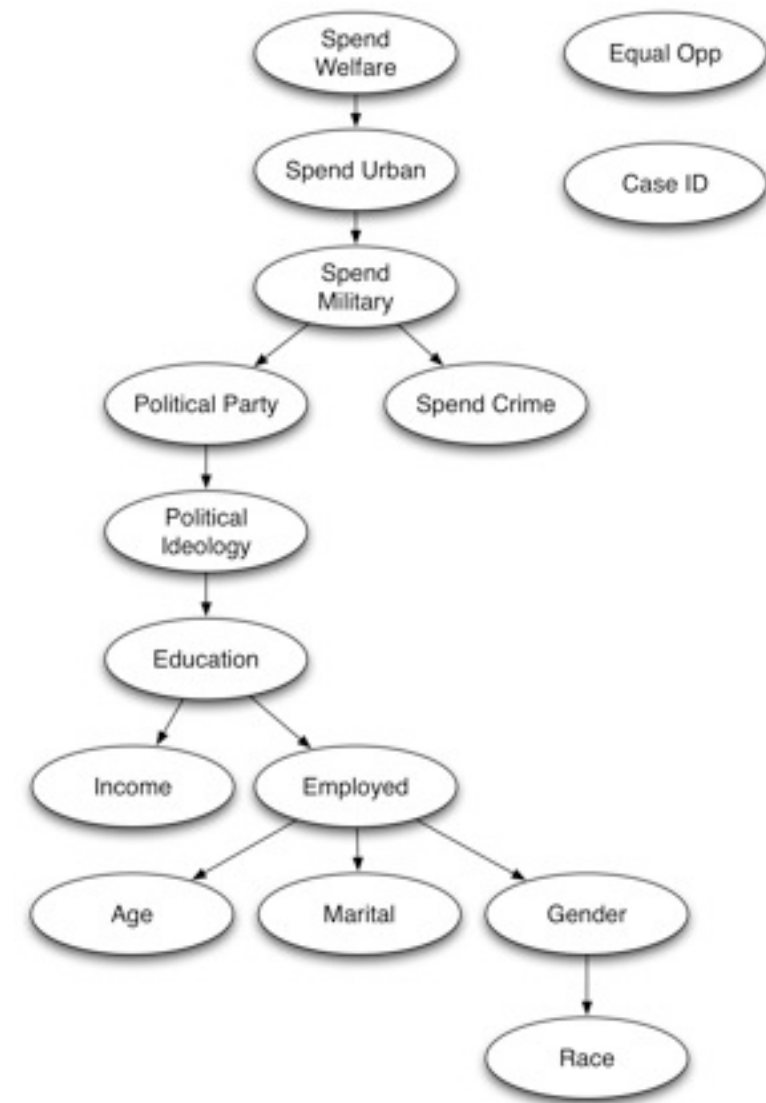
- Clinical protocols, medical records
- Human rights monitoring
- Agricultural certification and control
- Supply chain, etc.
- Still, how do we design good forms?



Usher

Data-driven Form Design

- Learn a probabilistic model
- Inputs: Form and data
- Outputs: Distributions and relationships
- Use model to improve form layout and presentation



Kuang Chen



Automatic Form Layout

- Greedy algorithm asks “most unpredictable question first”, within specified constraints
- Dynamic for mobile or web surveys
- Better prediction of remaining fields
 - Adaptive widgets
 - Maximize information gain when respondent is interrupted, or *satisficing*

Adaptive Widgets

- *Reduce effort*: make it easier (and more likely) to enter high-probability values
- *Improve accuracy*: warn user when high probability of being wrong

6. * Subcounty

Kabuyanda ▼

- Nyakitunda
- Kabuyanda
- NA--
- Birere
- Kabuyanda
- Kikagati
- Mwizi
- Nyakitunda

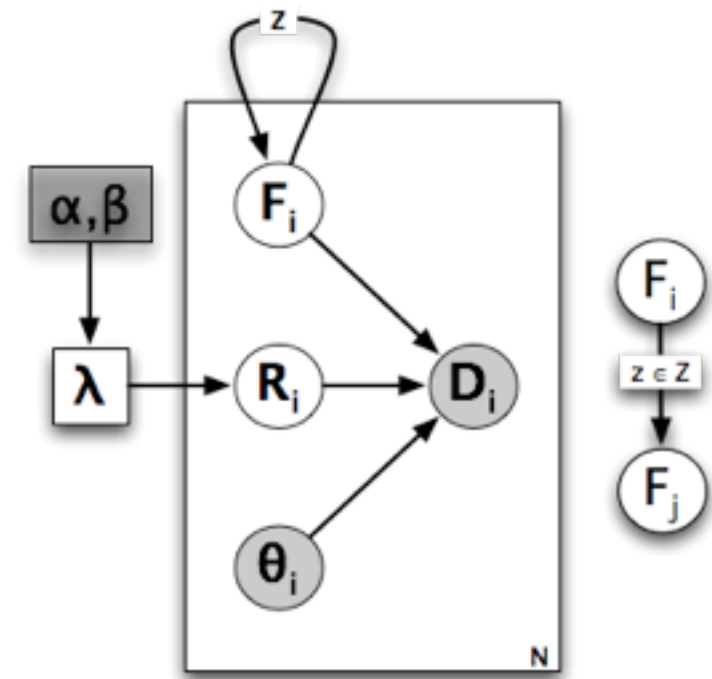
Family name *

As

- Aslimwe
- Aslimire
- Asasira
- Asingwire
- Asasire

Selective Double-Entry

- Double entry is gold standard for data quality
- Selectively re-ask those likely to be incorrect
- Immediately, or at end
- Tunable for cost / accuracy trade-off

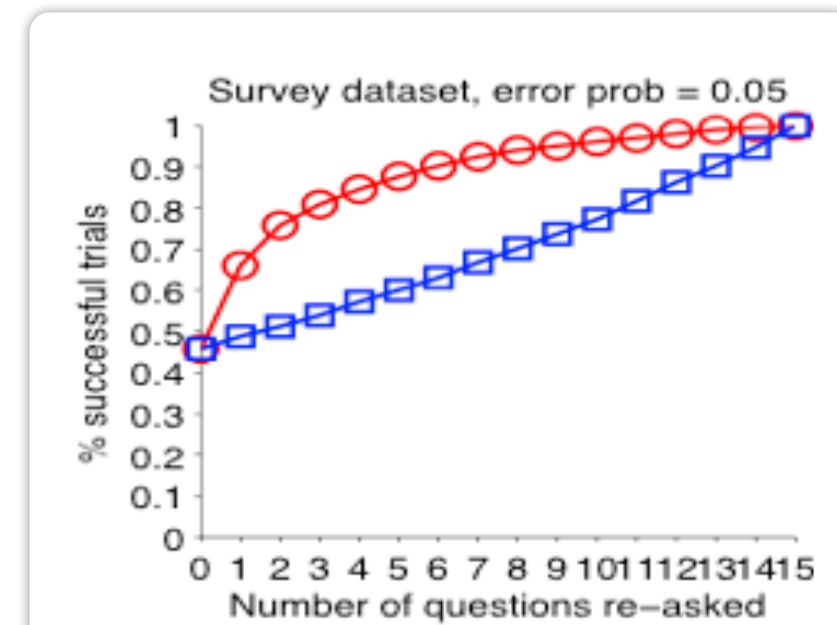


Error model

F_i : true unknown value
 D_i : noisy entered value
 λ : error distribution
 R_i : binary variable indicating whether there is an error
 Z : connects correlated fields

Current Status

- Simulation and user studies indicate potential to significantly increase data quality, and to reduce effort
- Packaging Usher as an open source tool that integrates with existing web and mobile software



Asking the right question

Choose your questions wisely

Asking the question right

Make it easier to be accurate

Summary

- Development is a process of iterative learning by doing, or *Design*
- IT not only provides new solutions; but can facilitate learning from experience
 - Avaaj Otalo
 - Usher

Big Challenges

- Empiricism at any scale
- Protecting, and leveraging, linguistic variation
- Searching user-generated multimedia
- Putting people first!

Thanks!

- Kaushik Ghosh, Apala Chavan, Sarit Arora, Puneet Syal, Neil Patel, Deepti Chittamuru, Paresh Dave, Anupam Jain, Nitendra Rajput, Amit Nanavati, Kuang Chen, Harr Chen, Neil Conway, Joseph Hellerstein
- CCD, DSC, Millennium Villages, Media Lab Asia, HFI, IBM Research India
- Nokia Research, Bill & Melinda Gates Foundation, Stanford, UC Berkeley

