

Intelligent User Interfaces

MICHAEL BERNSTEIN

CS 376

If you wanted a smart doorbell...

- To automatically control entrance to your room
- To let in possible donors for your Stanford education



No way



Let him in

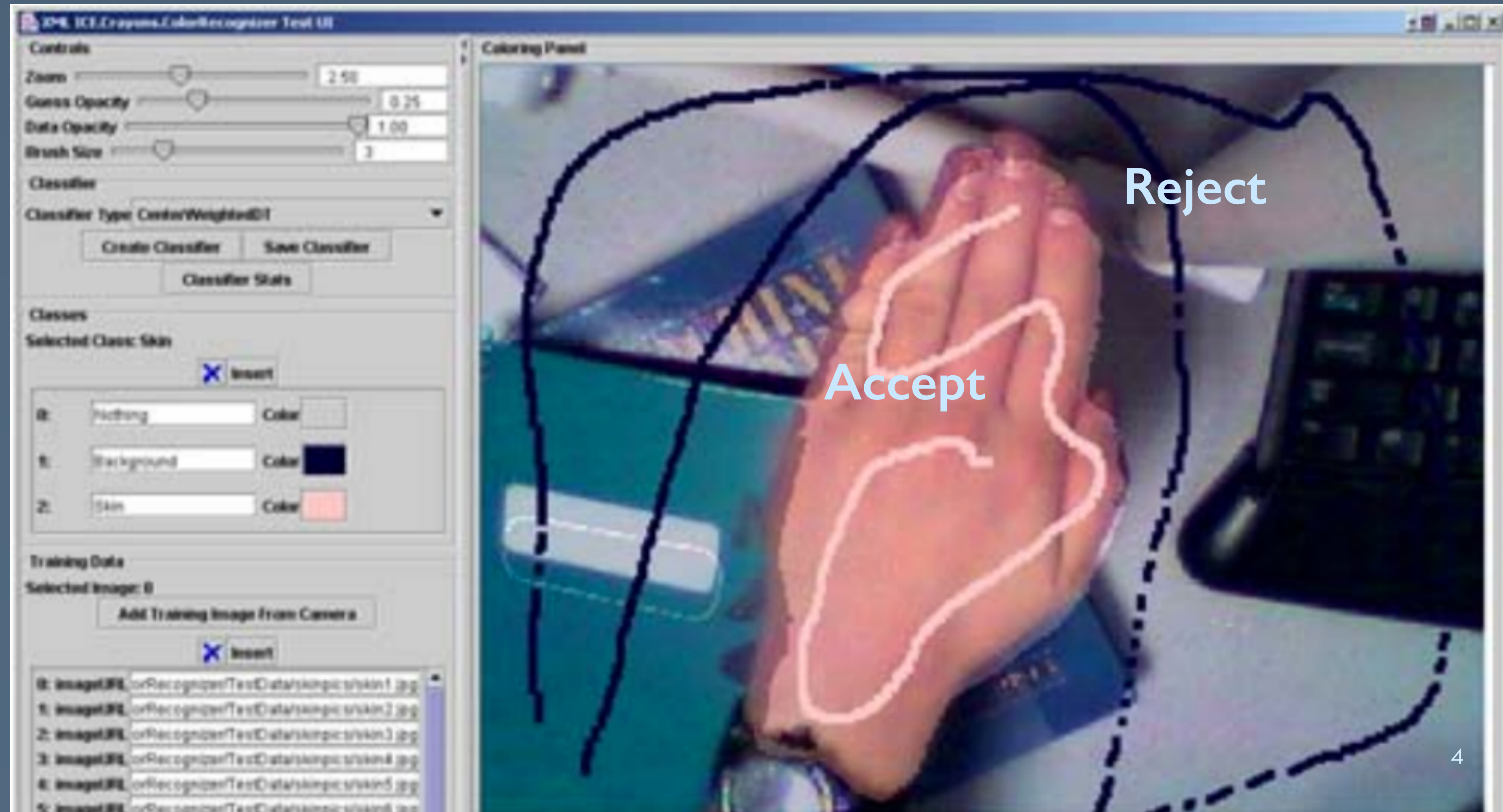
Paint the areas of interest



Crayons: camera-based interaction

[Fails and Olsen, CHI '03]

Direct-
manipulation
training



Intelligent user interfaces

- Goal: fashion powerful, easy-to-use interactive systems that draw on artificial intelligence and machine learning
- Challenges
 - Training
 - Development
 - Applications
 - User control

IUI applications

Recall: programming with screenshots

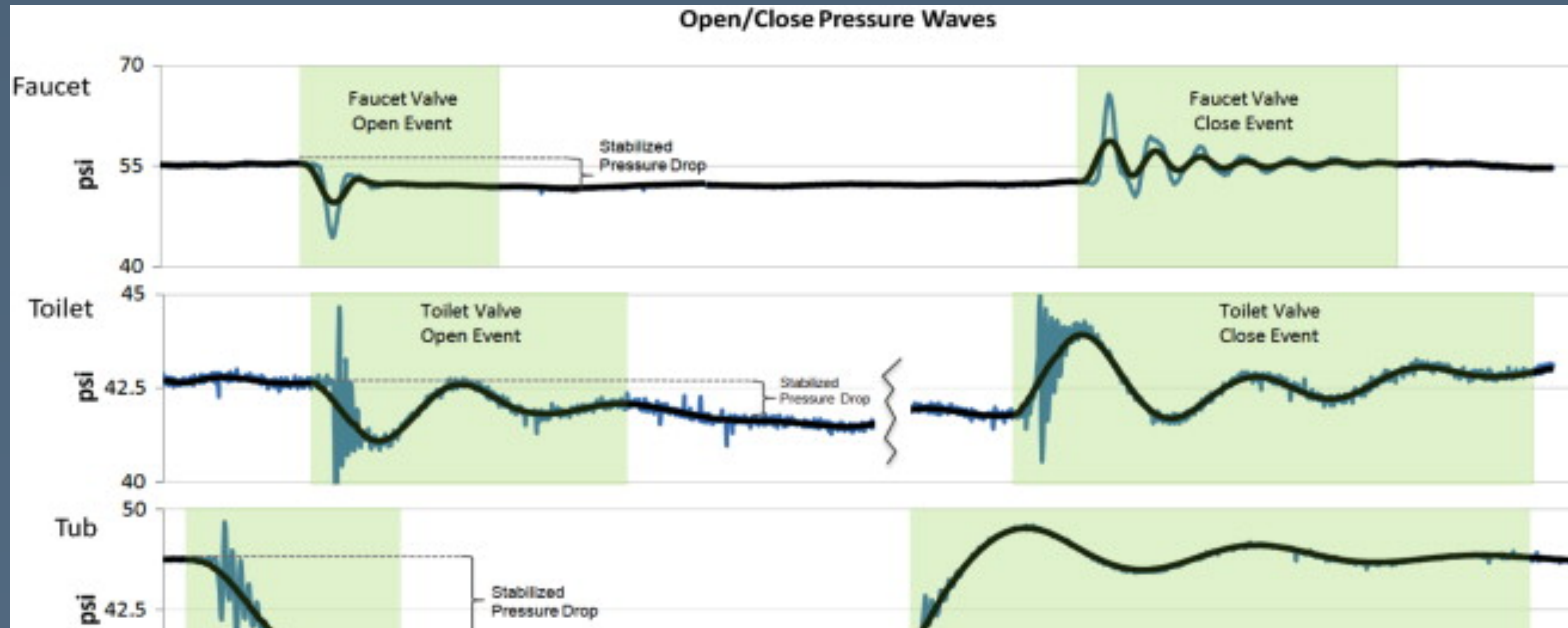
[Yeh, Chang, and Miller, UIST '09]

- Template search in desktop scripting



Recall: infrastructure-mediated sensing

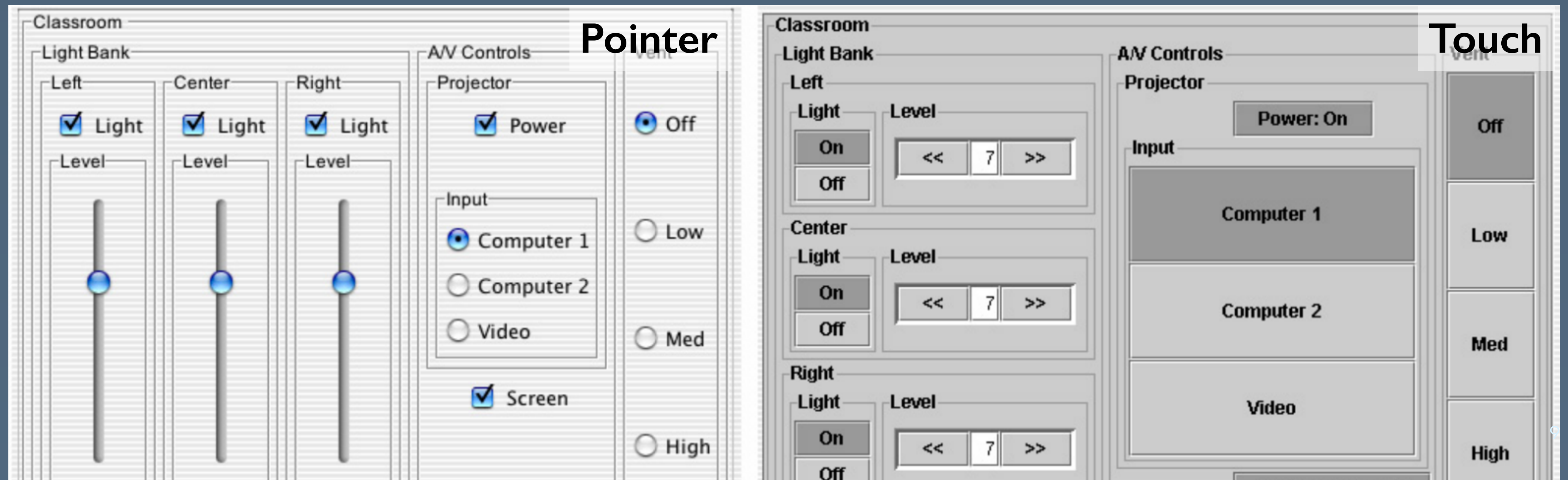
- Learning to classify in-home events



Automatically generating interfaces

[Gajos and Weld, IUI '04]

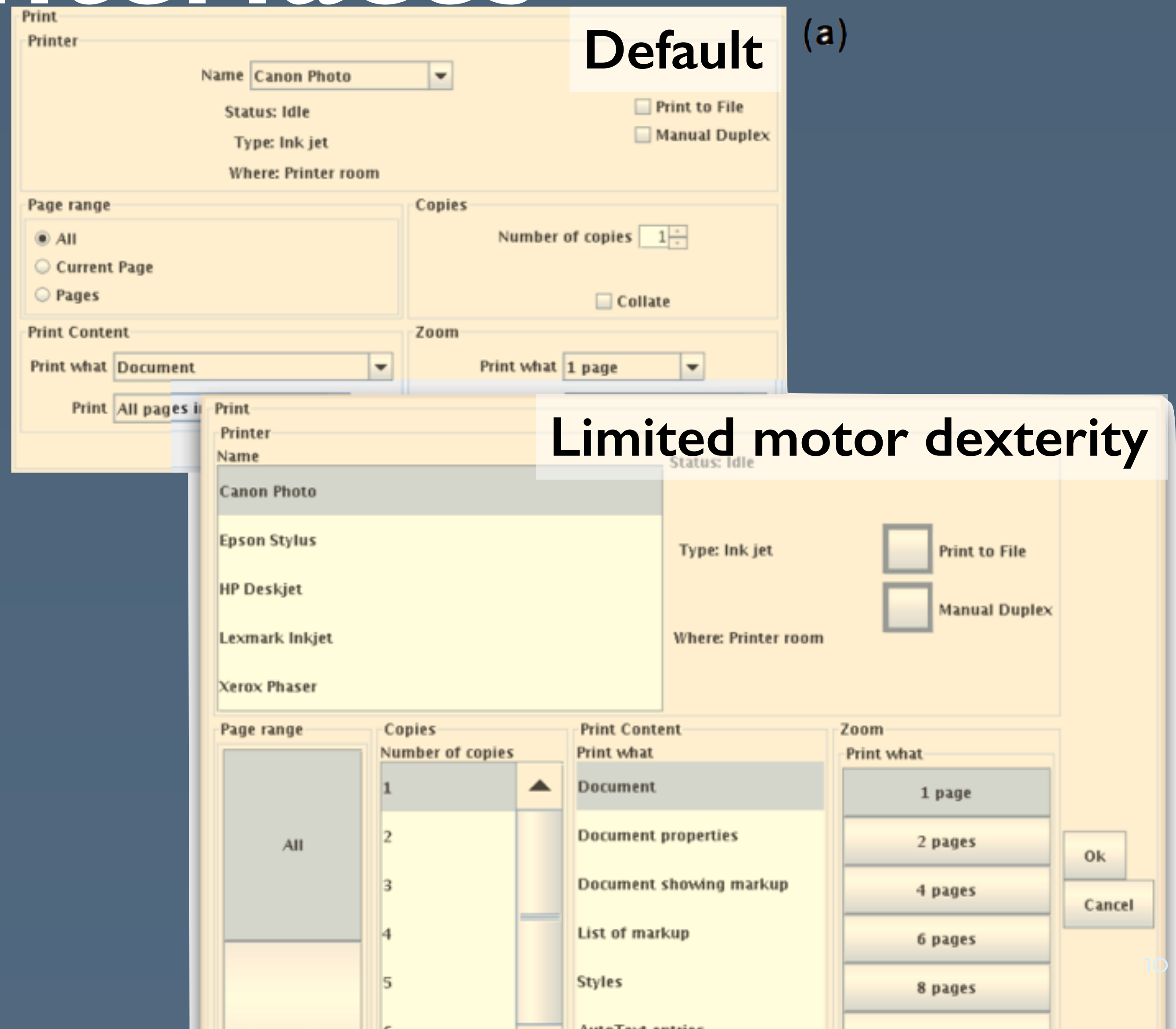
- Reactive design: remaps to output affordances
- Minimize a cost function derived from navigating between widgets in user traces



Ability-based interfaces

[Gajos et al., CHI '08]

- Rather than adjust to the device, adjust to the person
- Motor tests measure abilities of disabled individuals
- 25% faster, 73% fewer errors with automatic SUPPLE adjustment



Accelerating information extraction

[Hoffmann et al., CHI '09]

1. Automatically extract structure from text
2. Ask web site users to verify or correct

The image shows a screenshot of a web page with a yellow header that reads "Jerry Seinfeld". Below the header is a photograph of Jerry Seinfeld in a tuxedo. To the left of the photo is a snippet of text: "Jerome 'Jerry' Seinfeld (born on April 29, 1954 in Brooklyn, New York) is a Golden Globe and Emmy Award winning American comedian, actor and writer...". Two yellow pop-up boxes are overlaid on the page. The first pop-up, on the left, asks "Is this correct for Jerry Seinfeld?" and displays extracted information: "birth_place" with the value "Brooklyn, New York". The second pop-up, on the right, asks "Is this correct for Jerry Seinfeld?" and displays extracted information: "birth_date" with the value "April 29, 1954". Both pop-ups have "Yes" and "No" buttons at the bottom.

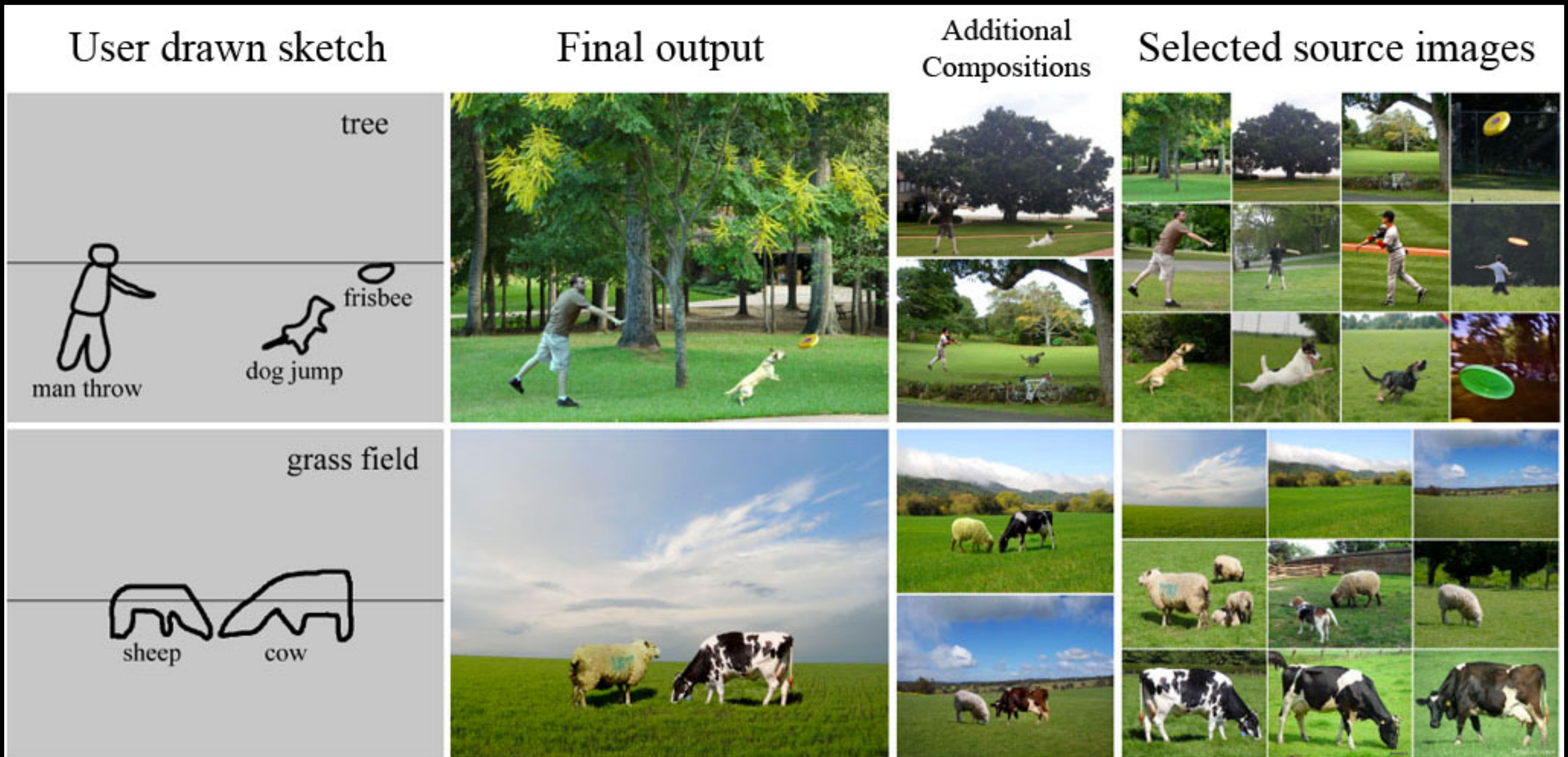
Completing sketched input

[Chen et al., SIGGRAPH Asia '09]



Completing sketched input

[Chen et al., SIGGRAPH Asia '09]

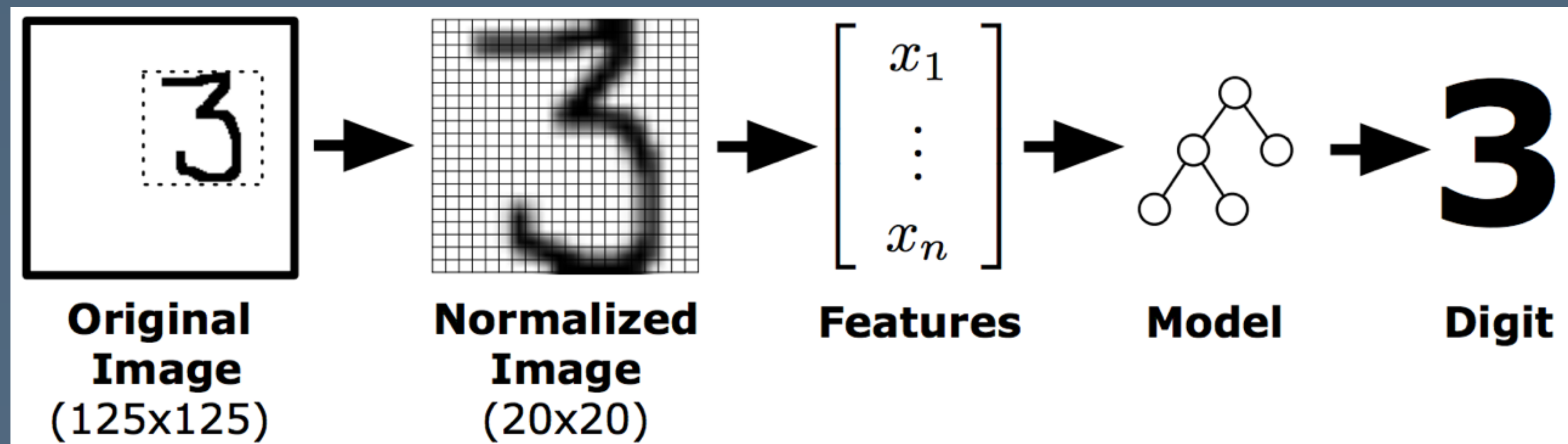


Developing intelligent software

Development challenges with ML

[Patel et al., CHI '08]

- Software development benefits from modularity, but machine learning is iterative and nonlinear
- Difficulty understanding the statistical process underlying machine learning algorithms
- Evaluation of progress is difficult



Papier-Mâché: toolkit support for tangible input

[Klemmer et al., CHI '04]

- Monitoring window, wizard-of-oz input, of-oz input, listeners, designed and evaluated as a user interface

The screenshot displays the Papier-Mâché software interface, which is divided into several panels:





- Current Phobs:** A hierarchical tree view showing the system's state. It includes folders for RFIDReaders, VisionPhobGenerators, and BarcodePhobGenerators. Under VisionPhobGenerators, there is a sub-entry for "VisionPhobGenerator: (Video Camera)" with five colored circular markers representing detected objects. Each marker is associated with a set of bounding box coordinates, size, hue, and saturation values. For example, a red marker has "Bounds[148, 7, 219, 124]; Size[1753]; Hue[196], Sat[94]".
- Vision input (Video Camera):** A video feed showing a wooden surface with a green sticky note, a red pen, a small white card with a barcode, and a pair of black scissors. A green dot is visible on the scissors' handle. A timestamp "00:00:46.65" is shown at the bottom of the video frame.
- Classifier Map:** A panel titled "Classifier Map" showing a "TagPhob" and "VisionPhobGenerator: (Video Cam)". Below this is a "New classifier" button.
- Association Map:** A panel titled "Association Map" displaying a list of bounding box coordinates and sizes for detected objects, such as "Bounds[22, 48, 63, 127]; Size[335]" and "Bounds[22, 48, 63, 127]; Size[336]".
- Vision Processing:** A panel titled "Vision Processing" with a sub-section "Edge Detection Thresholds". It features three sliders, each ranging from 0 to 255, used to adjust the sensitivity of the edge detection algorithm.

IDE support for ML development

[Patel et al., UIST '10]

- Explicit support for each step: feature extraction, model generation, training and testing

The screenshot displays an IDE interface with a legend on the left and a classification results table on the right. The legend lists various classes with corresponding colored squares, such as 'arrow (330)', 'caret (330)', 'check (330)', 'circle (330)', 'delete_mark (330)', 'left_curly_brace (330)', 'left_sq_bracket (330)', 'pigtail (330)', 'question_mark (330)', 'rectangle (330)', 'right_curly_brace (330)', 'right_sq_bracket (330)', 'star (330)', 'triangle (330)', 'v (330)', and 'x (330)'. The classification results table shows two rows of data, each with a 'Stroke' and a 'Normalized Stroke' image, and columns for 'Class', 'RowNumber', 'Number', 'Milliseconds', 'Name', 'Subject', 'AppName', 'AppVer', 'TimeOfDay', and 'Date'.

Raw Object	Class	RowNumber	Number	Milliseconds	Name	Subject	AppName	AppVer	TimeOfDay	Date
 Stroke	arrow	0	1	547	arrow01	1	Gestures	3.5.0.0	5:05:00 PM	Monda
 Normalized Stroke										
 Stroke	arrow	1	2	557	arrow02	1	Gestures	3.5.0.0	5:05:01 PM	Monda
 Normalized Stroke										

Play-along learning

[Fiebrink, Cook, and Trueman, ICMC '09]

- Create the output (sounds) you desire
- “Play along” and demonstrate the input that should generate that output



Who's in charge?

Software agents

- Delegate to proactive software and artificial intelligence

Pattie Maes, MIT Media Lab



Direct manipulation

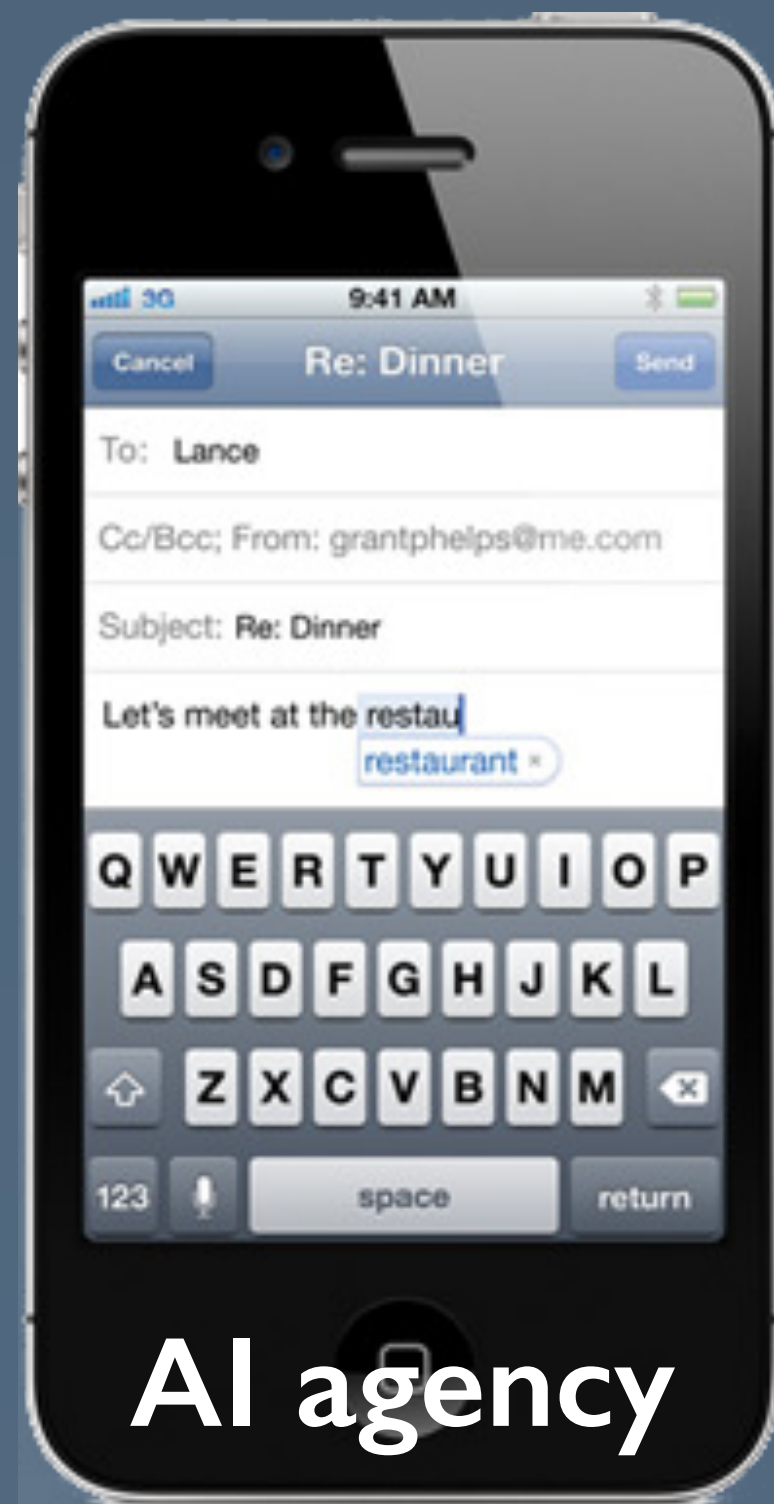
- Users should always have full control

Ben Shneiderman, U. Maryland



Mixed-initiative interaction

- Software proposes, user decides
- Removes the risk that the system may be incorrect, reduces user effort



Mixed
initiative

human-computer interaction

human-computer interaction

human-computer interaction degree

human-computer interaction journal

human-computer

Multimodal interaction

Using simultaneous inputs

- Sensor fusion can help disambiguate multiple noisy signals



Put That There

Speech N-best	Gesture N-best	Multimodal N-best
Zoom in	Checkmark	Zoom out
Show info		
Show all		
Zoom out		

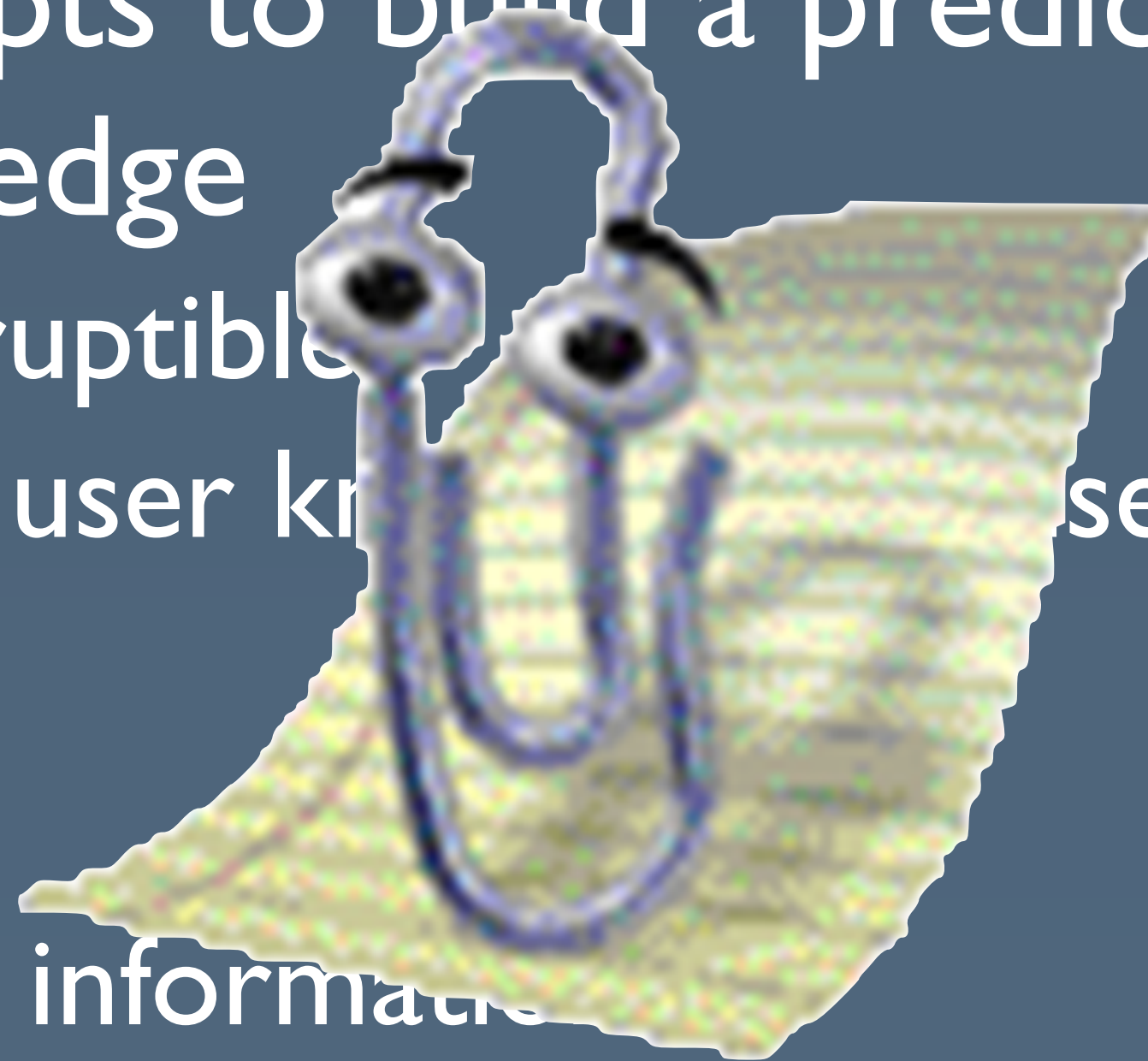
Quickset

[Oviatt, CHI '99]

User modeling

Software that knows you

- User modeling attempts to build a predictive model of the user's state or knowledge
 - State: is the user interruptible?
 - Knowledge: would the user know how to use this tool in Photoshop already?
- Challenges
 - Where does it get this information?
 - What if it's wrong?



IUI research: next steps

- Opportunities
 - IUI research can drive new insights in machine learning research
 - Machine learning skills may enable interactive systems that others cannot envision
- Challenges
 - System behavior can be unpredictable
 - Difficult to build user trust