

05- Methods and Experiments

Lorrie Cranor, Blase Ur,
and Rich Shay

January 27, 2015

05-436 / 05-836 / 08-534 / 08-734

Usable Privacy and Security



Today's class

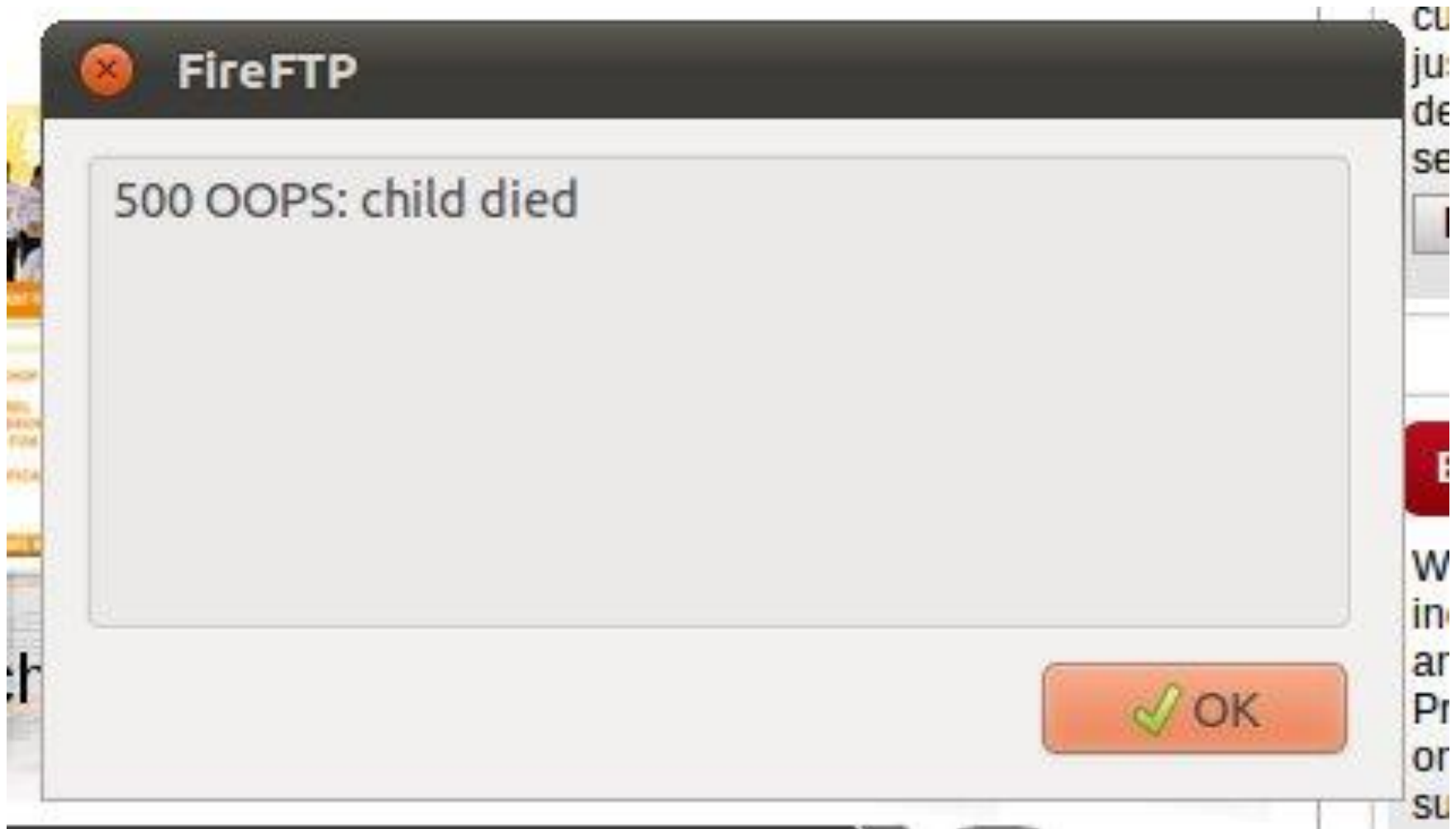
- General HCI design methods
- Types of research studies
- Overview of research methods
- Study logistics and validity
- Participant recruitment
- Deception and ethics

Human-Computer Interaction (HCI)

- You are not the user! You know too much!
- Think about the user throughout design
- Involve the user



Human-Computer Interaction (HCI)



What is usable?

- Intuitive / obvious
- Efficient
- Learnable
- Memorable
- Few errors
- Not annoying
- Status transparent

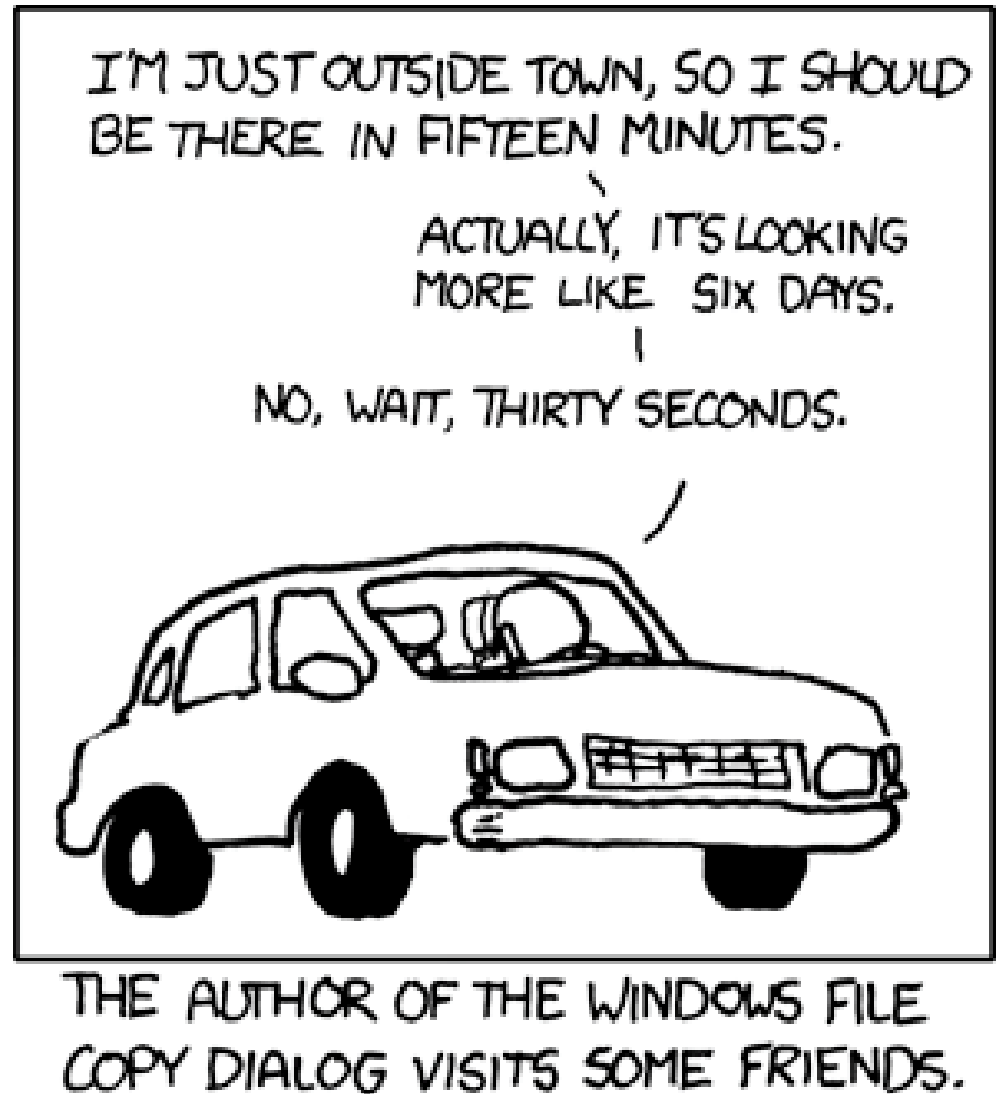


Image from <http://www.xkcd.com>

Difficulties

- Many systems and platforms
- Users are different from one another
- Required standards (or no standards)
- Documentation won't necessarily be read
- Performance
- Legal / time pressures
- Social and external factors

Determine use cases and goals

- What are the concrete tasks users should be able to accomplish?
 - Based on understanding of users!
- Set realistic metrics

Personas (example)



Name: Patricia

Age: 31

Occupation: Sales Manager, IKEA Store

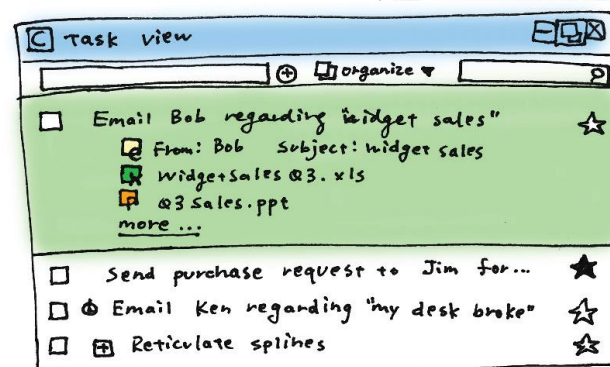
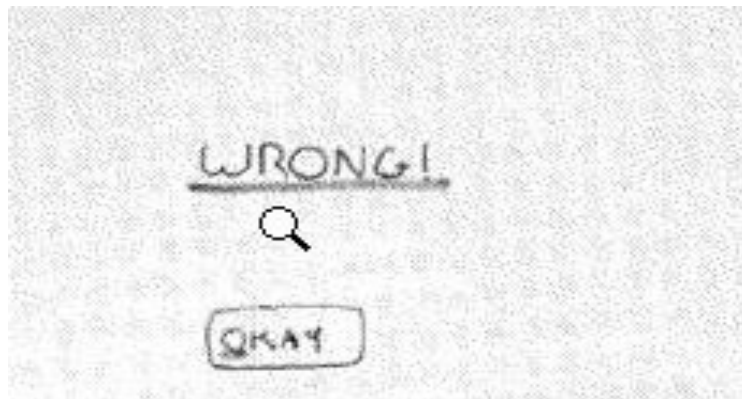
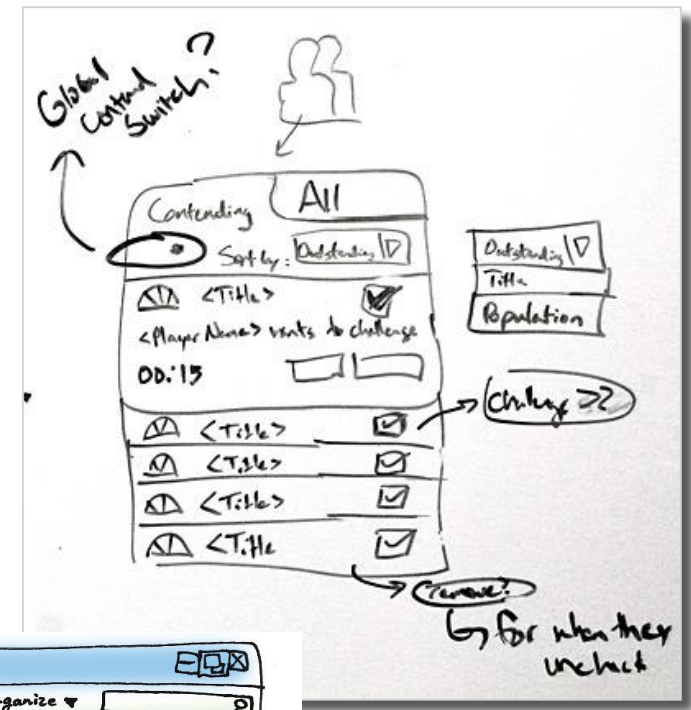
Hobbies: Painting
Fitness/biking
Taking son Devon to the park

Likes: Emailing friends & family
Surprises for her husband
Talking on cell phone with friends
Top 40 radio stations
Eating Thai food
Going to sleep late

Dislikes: Slow service at checkout lines
Smokers

Iterative prototyping is crucial!

High-fidelity, "Wizard of Oz," low-fidelity



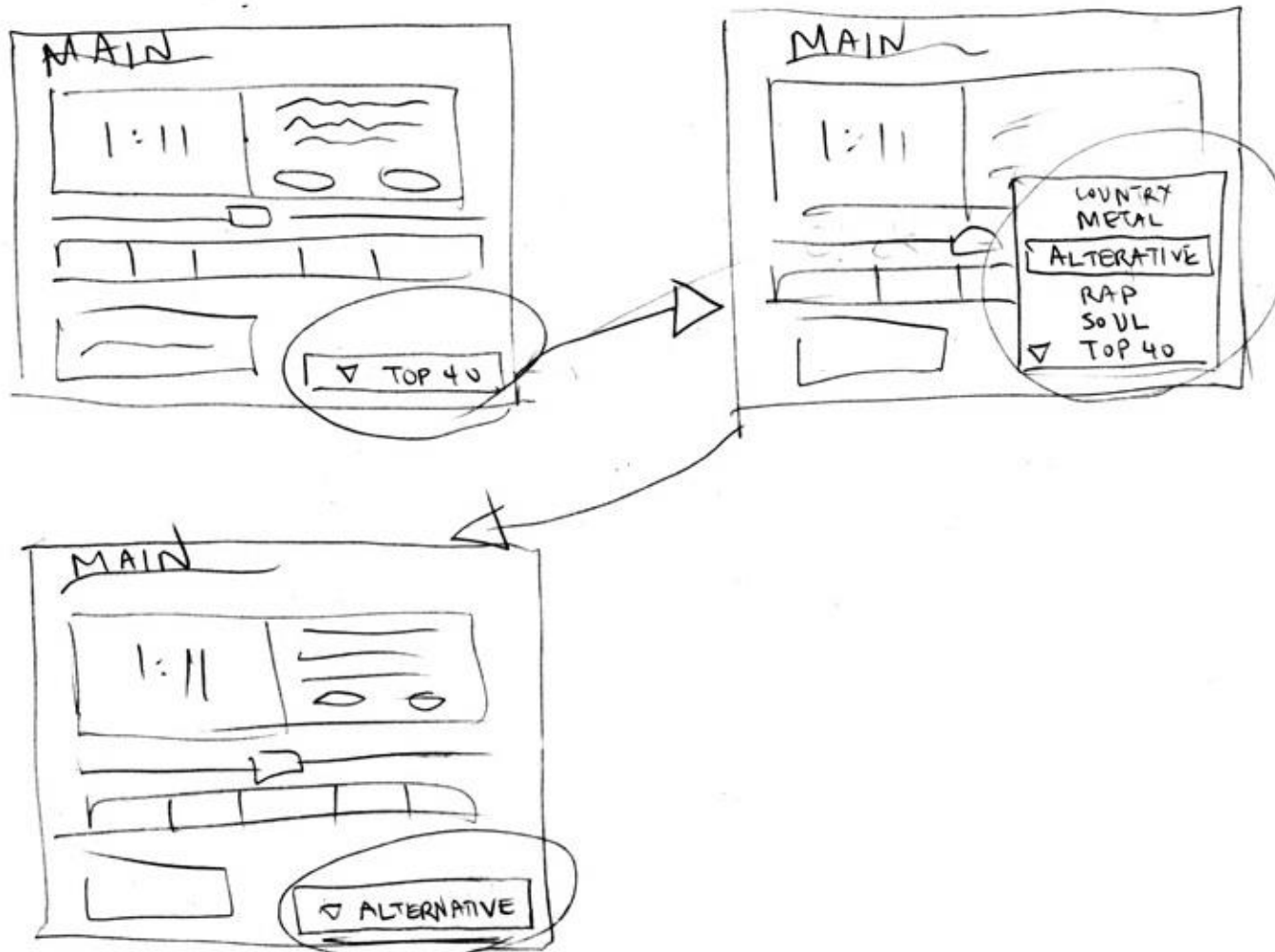
Paper prototypes

- Don't overthink. Just make it.
- Draw a frame on a piece of paper
- Sketch anything that appears on a card
- Make all menus, etc.
- Redesign based on feedback
- “Think aloud”

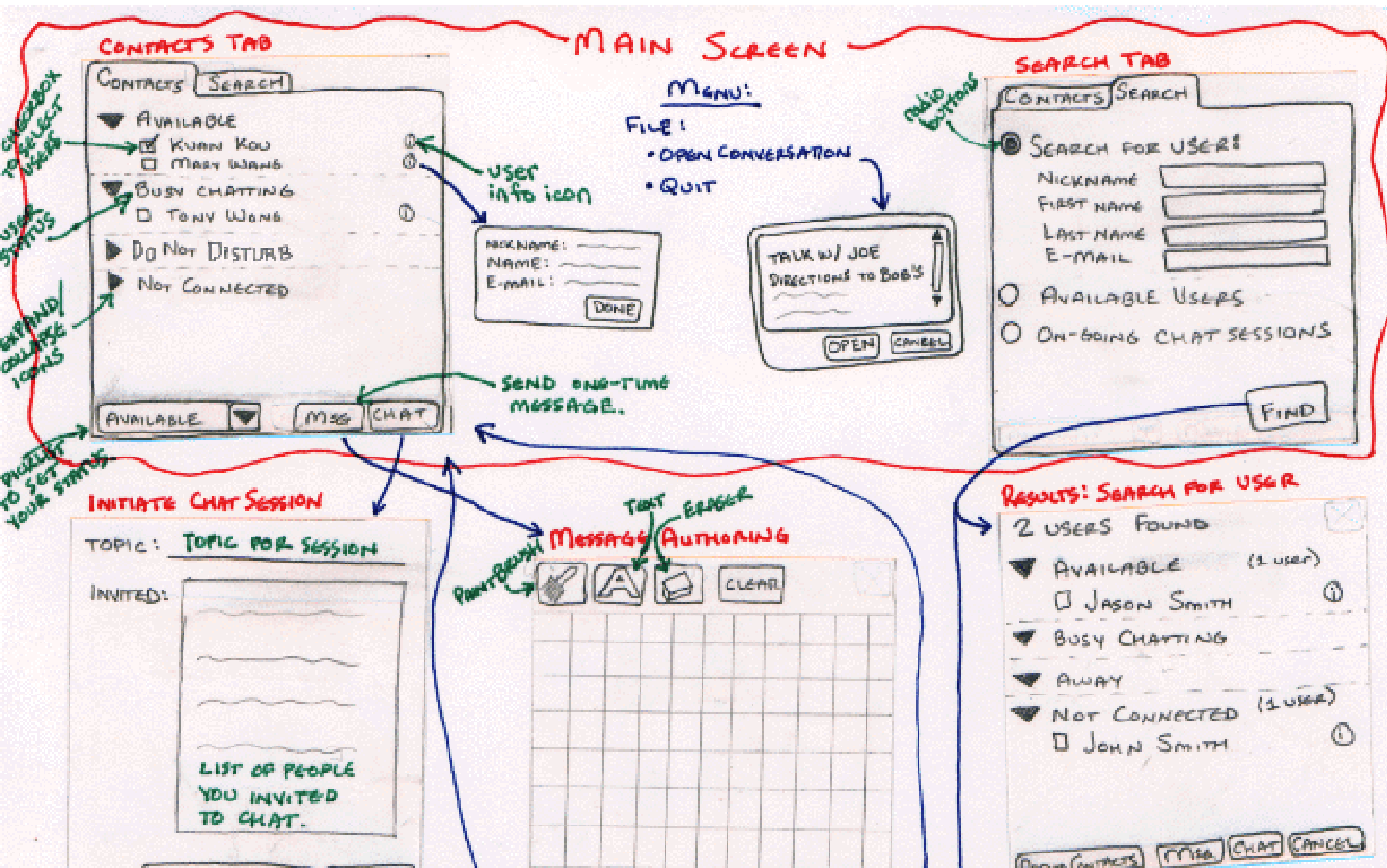
Paper prototypes

SCENARIO 1

"I want to listen to alternative music"

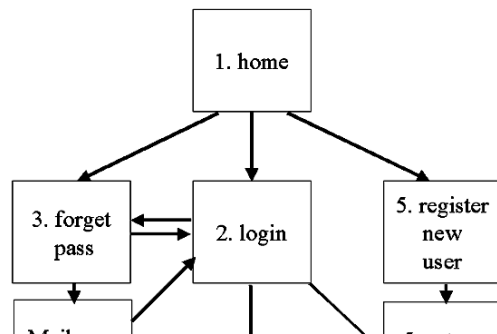


Iterative prototyping is crucial!

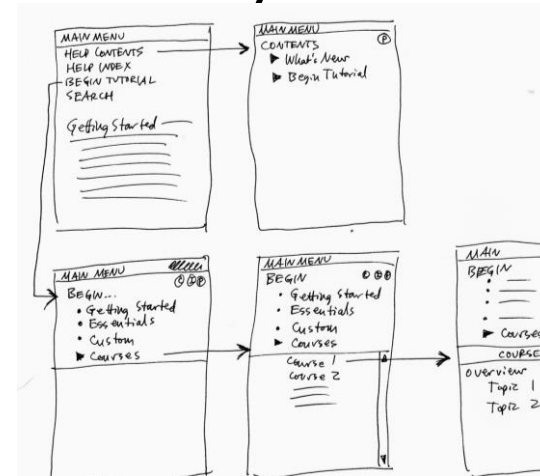


Usability prototyping for websites

Site Maps



Storyboards



Schematics

Sales Home	(Site Branding)	
Acme, Inc. Kids Outdoors Catalogue Travel Features About This Site (global nav bar)	(What this site is about) Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat.	Contact email Search: width = x char
	News Topic	
	This month's news release (date) This month's news release (date)	
	News Topic	
	This month's news release (date)	
	News Topic	
	This month's news release (date)	
	News Topic	
	This month's news release (date)	
	Acme, Inc. - sales home	
<u>section 1 - section 2 - section 3 - section 4 - section 5</u> <u>section 6 - section 7 - section 8 - section 9</u>		

Mock-ups

Think aloud example

- Download and install software that lets you encrypt your email
 - “Think aloud” of whatever’s on your mind
 - Give them an example
- Additional things you can ask:
 - What are you thinking now?
 - What do you expect to happen if you do X?
 - How did you decide to do that?

Research studies: purpose and goals

- What are you hoping to learn?
- What are your hypotheses?
 - Sometimes listed explicitly in a paper
- What are your metrics for success?
 - More secure, quicker to use, more fun, etc.
- What are you comparing to?
- What data might be helpful?

Broad types of studies

- Descriptive study
 - Relational study
 - Experimental study
- 
- The image is a meme featuring a stick figure. The figure is holding a flask in its right hand, from which small circles representing bubbles are rising. In its left hand, it holds a smartphone. Above the figure, the words "STAND BACK" are written in a large, bold, black, sans-serif font. Below the figure, the words "I'M GOING TO TRY" are written in a smaller, bold, black, sans-serif font, followed by the word "SCIENCE" in a very large, bold, black, sans-serif font.
- Formative (initial) vs. summative (validate)

Quantitative vs. Qualitative

- Quantitative: you have numbers (timing data, ratings of awesomeness)
- Qualitative: you have non-numerical data (thoughts, opinions, types of errors)

Types of studies

- What people want/think/do overall:
 - Surveys
 - Interviews
 - Focus groups
- What people want/think in context:
 - Contextual inquiry (interviews)
 - Diary study (prompt people)
 - Observations in the field

Types of studies

- Expert evaluation of usability:
 - Cognitive walkthrough
 - Heuristic evaluation
- Usability test:
 - Laboratory (“think aloud”)
 - Online study
 - Log analysis

Types of studies

- Controlled experiments to test causation
- Varying different conditions
 - Full-factorial design or not
 - Independent and dependent variables
- Many methods apply (e.g., surveys can be designed to test causation)
 - Role-playing studies
 - Field studies

Data to collect during experiments

- Performance (time, success rate, errors)
- Opinions and attitudes
- Actions and decisions
- Audio recording, screen capture, video, mouse movements, keystrokes

Even more data to collect

- Demographics
 - Age, gender, technical background, income, education, occupation, location, disabilities, first language, privacy attitudes, etc.
- Open-ended questions
- Preferences and attitudes

Please respond to the following statements:

**This user interface was difficult to understand*

1- Strongly disagree 2- Disagree 3- Neutral 4- Agree 5- Strongly agree

**This tool was fun to use*

1- Strongly disagree 2- Disagree 3- Neutral 4- Agree 5- Strongly agree

Logistics for a study

- How many participants?
 - Statistical power
 - Time, budget, participants' time
- What kind of participants?
 - Skills, background, interests
 - Their motivations
 - Often not a “representative sample”
- What do you need to build, if anything?
 - Prototype fidelity

Study designs

- Within subjects
 - Every participant tests everything
 - Crucial to randomize order! (learning effect)
 - Fewer participants
- Between subjects
 - Each participant tests 1 version of the system
 - You compare these groups
 - Groups should be similar (verify!)
 - Still randomize!

Validity

- Is this study **ecologically valid**?
 - Does it mirror real-life conditions and context?
- To what degree can we generalize about our results (**externally valid**)?
 - What biases does our sample introduce?

Participants, ethics, and deception

Participants

- Recruit people to do something remotely (e.g., online)
- Recruit people to come to your lab
- Recruit people to let you into their “context”
- Observe people (if possible, get consent! If not possible, consider necessity of design)

Participants

- What recruitment mechanisms?
 - Craigslist, flyers, participant pools, representative sample, standing on street
- How do you compensate them?
 - Ethics of paying \$0.00 vs. \$10.00 vs. \$100,000
- How do you get informed consent?
- What happens to their data?
- Prior knowledge / “what” are they?

Ethics

- How do we protect participants?
 - What risks do we introduce?
- Is there a less invasive method that would give equivalent insight?
- IRB is one arbiter of ethics; experimenters themselves are another crucial arbiter
- How do we make sure participation is voluntary throughout the experiment?

Deception

- Do we mind if participants know precisely what is being studied?
 - Sometimes, it's crucial that we observe their organic responses in context
- What “deception” or “distraction” task can we introduce?
- How do we maintain ethics?
- How do we debrief people at the end?

An entire university's passwords

- 25,000 faculty, staff, students at CMU
- What are their password characteristics?
- How guessable are their passwords?
- How do demographic factors correlate with password strength?
- How do these real passwords compare to leaked / collected passwords?



MAIN MENU

MY STORIES: 25

FORUMS

SUBSCRIBE

JOBS

RISK ASSESSMENT / SECURITY & HACKTIVISM

It's official: Computer scientists pick stronger passwords

Landmark study says people in business school choose weakest passwords.

by Dan Goodin - Nov 8 2013, 12:28pm EST

IDENTITY PRIVACY B4



Ethics questions

- How did we get people's passwords?
- How did we obtain consent?
- What ethical concerns are there?
 - What seemed to be done well?
 - What could have been done better?

Social phishing

- Use social networking sites to get information for targeted phishing
 - “In the study described here we simply harvested freely available acquaintance data by crawling social network Web sites.”
- “We launched an actual (but harmless) phishing attack targeting college students aged 18–24 years old.”

Social phishing

- Control group: message from stranger
- Experimental group: message from a friend
- Used university's sign-on service to verify passwords phished

Ethics

- How did they obtain consent?
- What ethical concerns are there?
 - What seemed to be done well?
 - What could have been done better?
- Who was potentially affected by the study?
- “The number of complaints made to the campus support center was also small (30 complaints, or 1.7% of the participants).”

Institutional Review Board (IRB)

IRB process

- Is it research? Are there human subjects?
- Full review vs. expedited vs. exempt
- Fill out and submit protocol
 - Include all study materials (e.g., surveys)
 - Include recruitment text and/or poster
 - Leave plenty of time