PRIVACY, SECURITY AND USABILITY

User Studies

INTRODUCTION TO USER STUDIES

Motivation for Designing User Studies

- Needfinding
  - What should we build?
    - Getting a better understanding of what people need
      - use those insights to create meaningful new products and services
  - Needfinding draws upon methods from anthropology, psychology, engineering and design planning
- Examine tradeoffs
  - Which approaches best fit particular needs?

Motivation for Designing User Studies

- Design Thinking
  - What are the key concepts and aspects for the products we are examining?
    - Typically used for development of new products

Need Finding/Design Thinking

- Need Finding

- Design Thinking

Need Finding

- Needfinding in BioDesign

Motivation for Designing User Studies

- Evaluate Design
  - Are requirements met?
- Gain a deeper understanding of the challenges
  - What are the main problems that need to be addressed?

Designing a User Study

- Identify research goals, questions, metrics, use cases
  - What are concrete tasks users should be able to accomplish?
    - Choose tasks users are likely to have in real life
      - E.g., “add a new client to the database”
    - Don’t lead participants
      - “click on top right button, then scroll down, etc.”
      - You want to test the usability of the product in real-life
  - What are realistic metrics?

Designing a User Study (cont.)

- Decide on type of study
  - User population to study, environment
- Design study protocol, create detailed study materials:
  - Scripts, surveys, scenarios, incentives, instrumentation, prototypes, recruiting materials, etc.
- Pilot and iterate on study design
  - Review original design, walkthrough with your partners
    - What should be improved?

Designing a User Study (cont.)

- Collect data
- Process and analyze results
- Repeat as needed
  - Some or all of these steps

Research Goals

- What are your study hypotheses?
- What are your metrics for success?
  - More secure, quicker to use, more fun, etc.
- What is the starting point?
  - Is there an existing tool that achieves similar goals?
  - Are there other tools/prototypes under development?
  - Is there another way to perform the same tasks without the new tool?

Broad types of studies

- Descriptive (observational) study:
  - Determine characteristics of population
  - Example: How concerned are people about privacy?
- Relational study:
  - studies relationship between two or more variables
  - Example: Are people who have been victims of computer viruses more likely to run backups regularly than people who have not been victims?

Broad types of studies

- Experimental study:
  - Artificially manipulates study factors
  - Goals is to determine effect of a certain treatment, condition, or intervention
  - Example: Are people assigned to use a password manager more likely to choose strong passwords
    - than participants assigned to remember passwords without a password tool?

System Evaluation Methods

- Analytic evaluation
  - Through system analysis
- Empirical evaluation
  - User observation and measurements are conducted

System Evaluation Methods

- Qualitative
  - Using words, pictures to evaluate design
    - Open ended questions
    - Get detailed information from a small number of users and specific context
  - Examine user thoughts, opinions, product perception, decision making factors

System Evaluation Methods

- Quantitative
  - Using numeric values to evaluate design
    - Compare data from a larger number of users
    - Timing results, number of errors, user ratings of product (on a scale)

Value-Action Gap

- People genuinely don't do what they say they would do
- Commonly referred to as **what people say vs. what people do**

Value-Action Gap

More than 60% of participants said they were "likely" or "very likely" to buy a kitchen appliance in the next 3 months. ***8 months later, only 12% had.*** *How Customers Think*, Gerald Zaltman, 2003

https://careerfoundry.com/en/blog/ux-design/the-importance-of-user-research-and-how-to-do-it/

Value-Action Gap Examples

- how much people say they care about recycling, and how much they recycle in practice.
- how many people say they favor organic food, and how many purchase organic food in practice.
- how much people say they care about carbon emissions, and how much they are willing to change their habits to reduce their carbon emissions in practice.

Value-Action Gap

- it's good to have a mix of both qualitative and quantitative data to draw from
  - so you don't run into issues from the Value-action gap
    - which can at times make qualitative data unreliable

System Evaluation Methods

- Formative (exploratory) studies
  - Usually conducted at the beginning of design process
    - Used to evaluate and refine ideas
  - Help provide insights into requirements, initial prototypes
  - Mostly qualitative, usually small sample size

System Evaluation Methods

- Summative (confirmatory) studies
  - Usually conducted at the end of design process
    - Test and evaluate the system
  - Evaluates success of design, tests hypotheses
  - Mostly quantitative, usually large sample size

Qualitative Methods

- Guerrilla testing:
  - fast and low cost testing methods
  - E.g., street videos, field observations, reviews of paper sketches or online tools for remote usability testing.
- Interviews:
  - one on one interviews that follow a preset selection of questions
  - prompting the user to describe their interactions, thoughts and feelings in relation to a product or service
    - or even the environment of the product/service.

Qualitative Methods

- Focus groups:
  - Participatory groups that are led through a discussion and activities
    - to gather data on a particular product or service.
- Field Studies:
  - heading into the user's environment and observing while taking notes
    - and photographs or videos if possible

Qualitative Methods

- In-Lab testing:
  - observations of users completing particular tasks in a controlled environment.
  - Users are often asked to describe out loud their actions, thoughts and feelings and are videoed for later analysis
- Card sorting:
  - Used to help understand Information Architecture and naming conventions better
  - Can be handy to sort large amounts of content into logical groupings for users

Quantitative Methods

- User Surveys:
  - Questionnaires with a structured format, targeting your specific user personas
  - These can be a great way to get a large amount of data. Surveymonkey is a popular online tool.
- First Click Testing:
  - A test setup to analyze what a user would click on first in order to complete their intended task.
  - This can be done with paper prototypes, interactive wireframes or an existing website.

Quantitative Methods

- Eye Tracking:
  - Measures the gaze of the eye, allowing the observer to 'see' what the user sees
  - This can be an expensive test and heatmapping is a good cheaper alternative.
- Heatmapping:
  - Visual mapping of data showing how users click and scroll through your prototype or website
  - The most well known online tool to integrate would be Crazyegg.

Quantitative Methods

- Web analytics:
  - Data that is gathered from a website or prototype it is integrated with
    - See demographics of users, pageviews and funnels of how users move through your site and where they drop off.
  - The most well known online tool to integrate would be Google Analytics.

Quantitative Methods

- A/B testing:
  - Comparing two version of a web page to see which one users prefer
  - This is a great way to test button placements, colours, banners and other elements in your UI.

UX Research Methods

- Research Methods Choice

Study designs

- Between subjects vs. within subjects design
- Within subjects:
  - Every participant tests all interfaces
- Between subjects:
  - Each participant experiences only one interface

Within Subjects Design Study

- Every participant tests all interfaces
- Much more powerful statistically
  - More data to analyze
- Measure the difference in how they do on the different interfaces
- Order of tasks needs to be randomized
  - Learning effect

Within Subjects Design Study

- Fewer participants
- Example:
  - Participants complete exercise tasks using both an Apple watch and Fitbit.

Between Subjects Design Study

- Each participant experiences only one interface
- Measure how well the people randomly assigned to the A interface did compared to the people randomly assigned to the B interface
- Requires more participants
- Still a good idea to randomize order of tasks
- Example:
  - Participants complete exercise tasks using either an Apple watch and Fitbit, but not both

Between Subjects Design Study

- Divide participants into separate group
  - Each participants in a group tests the same version of the system
    - Version different between groups
  - You compare these groups
  - Groups should be similar
    - Need to set criterions ahead of time for verification

Data to collect during experiments

- Performance results:
  - time, success rate, error rate
- Opinions, preferences, and attitudes
  - Does the product achieve its goals?
  - Is the user likely to use this product?
    - At what cost?
- Actions and decisions
  - Do decisions require certain knowledge or user background?

Data to collect during experiments

- Audio, screen capture, video, mouse movements, keystrokes
  - Do participants use the product in the way we anticipated?
- Demographics
  - Age, gender, technical background, income, education, occupation, location, disabilities, first language, privacy attitudes, etc.
    - How do demographics affect usability study results?

Data to collect during experiments

- Open-ended questions
  - Provide participants opportunity to point out important observations
    - Different users may have different insights into new product
    - Suggested Improvements?

Choosing Study Participants

- How many participants?
  - More participants, easier to run statistics
  - Constraints: budget, time, participants time
- What kind of participants?
  - Demographics:
    - Age, gender, etc
      - Important to note sample demographics in research documentation
    - Can we choose just students?
  - Skills, background, interests
    - That are relevant to the study
  - Frequently will not results in a representative sample

Study Validity

- Ecological validity
  - Extent to which the study findings can be generalized to real-life settings
    - How much study mirrors real-life conditions, context
  - Typically a trade-off with experimental control
    - Experimental control involves changing conditions under which the experiment occurs
      - These changes are different from what we would find in a natural setting
- External validity
  - The extent to which the results of a study can be generalized to other situations and to other people

SECURITY AND PRIVACY USER STUDIES

Security and privacy user studies

- What may you have in such a study?
  - Certain risks
    - Testing decision making under risk
  - Presence of an adversary

Testing conditions

- Test security and usability of systems when:
  - An attacker exists in the system
    - May try to deceive users
    - May change system functionality
      - Without the user knowledge
  - Users behave in unpredictable ways
    - Are the system still secure?
  - Users may be stressed or busy
    - Cognitive load may affect decision making
      - User may be more likely to make the wrong choice

Challenges

- Study should mimic real-life usage
  - Ecological validity
- Observing infrequent events
  - May not come up during the study
- Detect small differences between ways users perform tasks
  - Need careful assessment, video of participants
- Legal, ethical, practical issues
  - Plan a realistic study

Study options

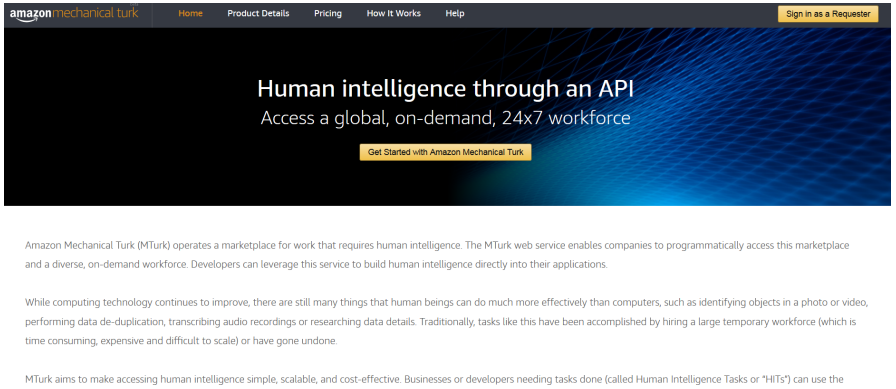
- Observe users ‘in the wild’
  - Data collection challenges
    - May create tool to run in the background of the device, gather user input
  - Hard to design a controlled experiment
  - Events of interest may be infrequent

Ask participants about hypothetical risks

- Design a survey
- Get participants engaged in the scenario
  - Through visualization, training
  - Simulate risks
- Ask participants about hypothetical scenario
  - What would they do in some cases
- Ask participants if they were subject to these risk in the past
  - What did they do

Large-scale online experiments

- Amazon Mturk for easy recruitment and payment
- Recruit, email, and pay participants without collecting personally identifiable information



Amazon MTurk

- Advantages:
  - Easy to execute
  - Relatively inexpensive
- Disadvantages:
  - Participants do it as a job
  - May rush, choose response randomly, not read questions, etc.
    - May not care about research integrity
  - Partial solution: add validation questions

Summary

- User-Centered design concentrates on user, tasks
- You are not the typical user!
- Usability studies help get quantitative and qualitative data from users
  - Help prevent mistakes early on
  - Test ideas
  - Test performance, timing, etc.

Questions?