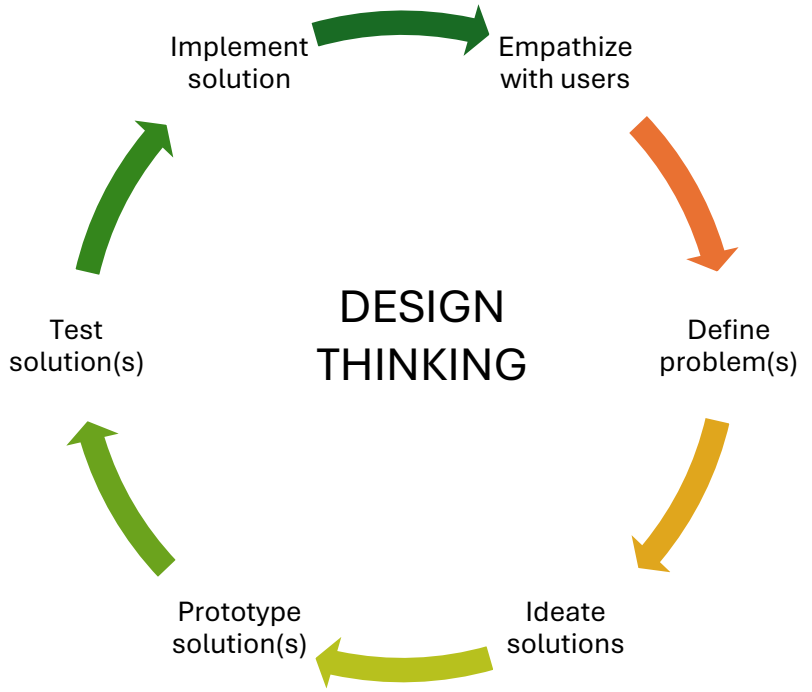
Lecture 14 – Evaluating prototypes (contd).

Last class

- Building and evaluating prototypes
- User studies as methods



How did the in-class evaluation go?

- Get feedback on Pingala prototypes?
- What worked, what didn’t work?

Conducting and analyzing usability studies

- Start with what you want to analyze usability for?
  - Entire tool vs. specific features?
  - If it is entire tool, break down to specific features (even multiple studies).

- Give users task(s) to do, that exercises the feature(s) of interest.
- Typically – more than one, of different kinds.
- For example, in ticket booking:
  - Find the best dates—you are planning a vacation at Dalhousie
  - Find the best tickets for a given date—you have a meeting in Dalhousie
  - Decide whether to book now or later, and go ahead with booking if you see fit (and stop just before entering personal details).

Common types of tasks to consider

- Use feature(s) to do a task the first time
  - E.g., add something
- Pick up a past task
- Make changes
- Delete stuff
- Find something, etc.
- Think of user flows, not features.

Task framing

- Frame in the users' task contexts and goals.
- People are not motivated if they have to "just do" something, for the sake of a study.
- Anchoring in real-world tasks helps them ground in actual behaviors.
- Also, important to get participants as close to real user(s), ideally from the user pool.

Participants

- Atleast 5, Usually 5-10.
- Smaller, if the user base itself is small.
- Get a variety, and from various circles
  - Email lists / posters
  - Prior databases
  - Purposive – for specific participant types
  - Convenience sampling –whoever you can find is okay, but not all your kind (ensure variety).
  - Snowball – ask participants for other participants.
  - Random / hallway – at random (whoever you see first!).
  - Stratified – to cover variety (e.g., for flights, this would look like students, business, leisure, men/women, people with children, etc.)

Conducting the study

- Almost always done one participant at a time
  - To not bias each other.
- Introduce participant to the tasks
- Show them the prototype
- Typically, no training – useful to see first time problems.
- Ask them to do the tasks
  - Think aloud as they do (helpful to see how they think)
  - Ask them to seek help

Data gathering: qualitative

- Record behaviors -- video + audio is useful
- Have someone make notes
- You can make notes yourself, if there are not too many tasks / questions, but always useful to have a second person / record.

Analyzing qualitative data

- You have what is called qualitative data
- Analyzing involves finding what is common among them
- “Thematic analysis”
  - Put finding into themes
- Examples:
  - Cant find search button on top right
  - Cant find sort on table header
  - Cant find filter options
  - => All point to “discoverability of common features” / “breaking conventions on feature placement” [One theme].
- List themes along with flavours and counts

Quantitative data

- Sometimes after tasks, it is also useful to get quantitative data
  - Likert scales (How did you like, how comfortable, etc.)
  - System usability scale (standard survey)
  - NASA Task load index (standard survey on cognitive load)
  - Feelings toolkits (people say how they felt using the tool –easy, convenient, frustrating, etc.)

- Analyzed in terms of means, medians, distributions, etc.

Pros and cons?

Cognitive walkthroughs

- Done within the team, when access to users is hard
- Even otherwise, do this as a first cut evaluation
- Pick a prototype and task; list the task steps
- Create a user “persona” (and list down their key characteristics)
  - Often, you need more than persona, then use ones at extremes
- For each step along the task, answer the following 4 questions
  - Will “User” want to do this? [Ideally, use persona name instead of “User”]
  - Assume “User” wants to do this, will s/he know what to do?
  - Assume “User” knows what to do, will s/he actually do it?
  - Assume “User” did it, will s/he know they did the right thing?
- Write down yes/no/maybe, along with reasons. Every no/maybe, is a usability issue to be fixed. The reasons often provide hints for what the fix is.
- Seems tedious, but can be done in an afternoon for atleast most common/least common paths as needed.

A note on personas

- Good personas are data driven, and come from user research (in the empathize phase)
- Example for how to do it, if you care:
  - https://uxpressia.com/blog/how-to-create-persona-guide-examples

- There are a lot of personas out there for use (people with disability, specific problem solving aspects, etc.)

- The definitive guide on the topic is:
  - “The persona lifecycle” by Tamara Adlin and John Pruitt.

Pros and cons?

Heuristic evaluation

- Heuristics = rules of thumb (for how to build interfaces)
- We evaluate interfaces against heuristics and look for violations
- Result is a list of violations to be fixed
- Who does it?
  - Ideally, someone that can interpret UI/UX heuristics, and catch violations (so a UI/UX expert/professional)
  - How many? Ideally, 5 or more.

- Provide interface (or screens)
- Provide a set of heuristics (or ask experts to pick their favorite)
- Evaluate interface against heuristics, and get a list of violations and possible fixes from the expert
- Challenging, but works great in large organizations with lots of designers, UX folks
- Also great for designers to evaluate their own designs systematically, as a first step

Example heuristics

- https://www.nngroup.com/articles/ten-usability-heuristics/
- Ben Shneiderman's eight golden rules of interface design
- Microsoft's guidelines for human-AI interaction design
- UI tenets and traps
- Rules from Steve Krug's "Don't make me think" for web usability
- There's a lot more, go look for them!

Pros and cons?

Pragmatic evaluations

- Pick and choose methods at various points in time
- Pick combinations
- Be creative!