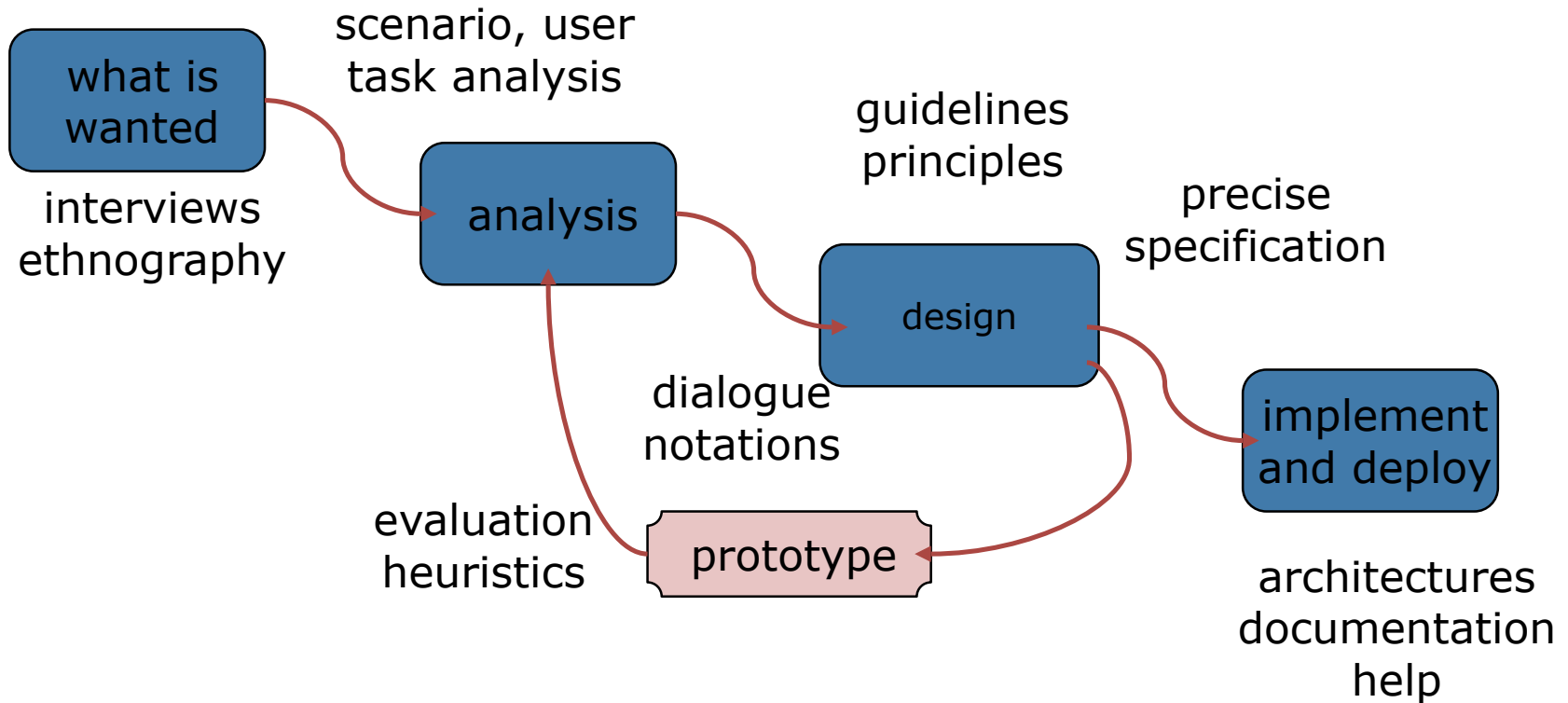


Human-Computer Interaction

Termin 7:
Usability Evaluation

User centered design

Process to develop interactive systems such that **usability** will be maximized.



Key questions for today

- How can the usability of a system be evaluated?
- How can usability problems be found?



Evaluation methods

- Basic Concept:
 - hardware vs. software
 - formative vs. summative
 - qualitative vs. quantitative
 - subjective vs. objective
 - analytic vs. heuristic vs. empirical
 - formal vs. informel
 - experimental vs. guided
 - theory-based vs. user-based



Evaluation methods

□ Inspection methods (Expert-Review)

- Guidelines review & consistency inspection
- Cognitive walkthrough
- Heuristic evaluation
- Focus group

□ Study of use

- Usability-Test
- Thinking-Aloud
- Field studies
- Interviews & questionnaires

□ Model-based Evaluation



Usability-Inspection methods

Guidelines Review
Consistency Inspection
Cognitive Walkthrough
Heuristic Evaluation
Focus Group

Guideline review

Consistency inspection

- System/interface is checked for conformance with guidelines
 - Standard guidelines,
 - Organization specific guidelines,

- Consistency inspection
 - of terminology, colors, fonts, icons, menus, general layouts, etc.
 - of interaction style



Usability-Heuristics (1)

- **Visibility of system status**
 - System should always keep users informed about what is going on
- **Match between system and the real world**
 - Should speak the users' language, with words, phrases and concepts familiar to the user. Follow real-world conventions, making information appear in a natural and logical order.
- **User control and freedom**
 - Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state (Support undo and redo)
- **Consistency and standards**
 - Users should not have to wonder whether different words, situations, or actions mean the same thing.



Usability-Heuristics (2)

- **Error prevention**
 - Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.
- **Recognition rather than recall**
 - Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another.
- **Flexibility and efficiency of use**
 - Accelerators --unseen by the novice user --often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.



Usability-Heuristics (3)

- **Aesthetic and minimalist design**
 - Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
- **Help users recognize, diagnose, and recover from errors**
 - Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
- **Help and documentation**
 - Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.



User studies

Thinking-Aloud
Cooperative Evaluation
Interviews & questionnaires
Usability-Test

User studies

- ❑ Study the interaction between **actual user** and system
- ❑ Involves measuring **representative users' performance** on **carefully prepared tasks** that are typical of the tasks, for which the system was designed
- ❑ Users testing can use
 - video and interaction logging to capture errors and frequencies and time of commands
 - think-aloud protocols
- ❑ User testing may be done in the lab or the field
- ❑ Users may be interviewed or complete questionnaires
 - Satisfaction questionnaires provide data about users' opinions



Lab studies

- Experiment under controlled conditions
 - specialist equipment available
 - uninterrupted environment

- Disadvantages:
 - lack of context
 - difficult to observe user cooperation

- Prevalent paradigm in psychology

Field studies

- Experiments dominated by group formation

- Field studies more realistic
 - *distributed cognition* ⇒ work studied in context
 - real action is *situated*
 - physical *and* social environment crucial

- sociology and anthropology – open study and rich data



Think Aloud

- User is observed while performing a *predefined* task
- User is asked to describe what...
 - he is thinking right now
 - he is expecting to happen
 - he is thinking is happening

- Advantages
 - simplicity - requires little expertise
 - can provide useful insight into user's mental model
 - can show how system is actually used
- Disadvantages
 - artificial test situation → cooperative evaluation
 - subjective and selective → several trials needed
 - act of describing may alter task performance



Cooperative Evaluation

- User evaluates together with expert, sees himself as collaborator in evaluation
- both can ask each other questions

- Additional advantages
 - less constrained and easier to use
 - user is encouraged to criticize system
 - clarification dialogues possible

- Problems with *both* techniques
 - generate a large volume of information (*protocols*)
 - 'Protocol analysis' crucial and time-consuming



Query techniques

□ Interviews:

- analyst questions user, usually based on prepared questions
- informal, subjective and relatively cheap
- can be varied to suit context, issues can be explored more fully, can unanticipated problems
- very subjective, time consuming

□ Questionnaires:

- Set of fixed questions given to users, need careful design!
- Style of questions: open vs. closed, scalar (judge a statement on a numeric scale), multiple-choice, ordering, negative vs. positive, ...
- Style of answers: text, ja/nein, number of options, ...
- reaches large user group, can be analyzed rigorously, less flexible, less probing



Surveys

- Written user surveys (usually online)

- Useful to gather important data about
 - user background (age, gender, education, etc.)
 - experiences with computers
 - job responsibilities

- Can use survey instruments to determine
 - personality style of user
 - familiarity with features
 - feeling state of using the system



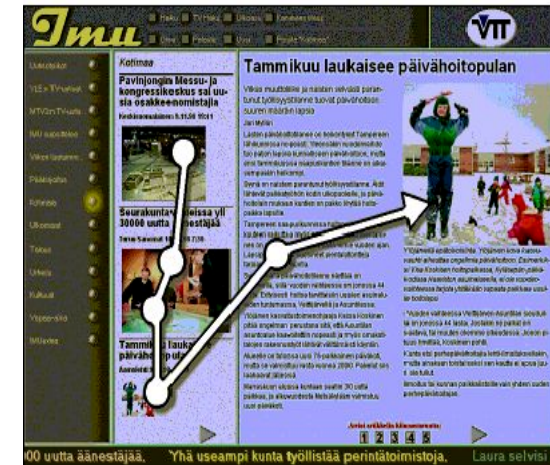
Physiological measurements

- Emotional response linked to physical changes
- may help determine a user's reaction to an interface
- measurements include:
 - heart activity, including blood pressure, volume and pulse.
 - activity of sweat glands: Galvanic Skin Response (GSR)
 - electrical activity in muscle: electromyogram (EMG)
 - electrical activity in brain: electroencephalogram (EEG)
- some difficulty in interpreting these physiological responses - more research needed



Eye tracking

- head or desk mounted equipment tracks the position of the eye
- eye movement reflects amount of cognitive processing a display requires
- measurements include
 - fixations: eye maintains stable position. Number and duration indicate level of difficulty with display
 - saccades: rapid eye movement from one point of interest to another
 - scan paths: moving straight to a target with a short fixation at the target is optimal



Recall: Methods in user-centered design

- Field studies (including contextual inquiry)
- User requirement analysis
- Iterative design
- Usability evaluation
- Task analysis
- Focus groups
- Formal heuristic evaluation
- User interviews
- Surveys
- ...

Ranking based on survey among experienced UCD practitioners
(103 questionnaires) (Mao et al., 2005)



Summary - Choosing an Evaluation Method

when in process:	design vs. implementation
style of evaluation:	laboratory vs. field
how objective:	subjective vs. objective
type of measures:	qualitative vs. quantitative
level of information:	high level vs. low level
level of interference:	obtrusive vs. unobtrusive
resources available:	time, subjects, equipment, expertise

