

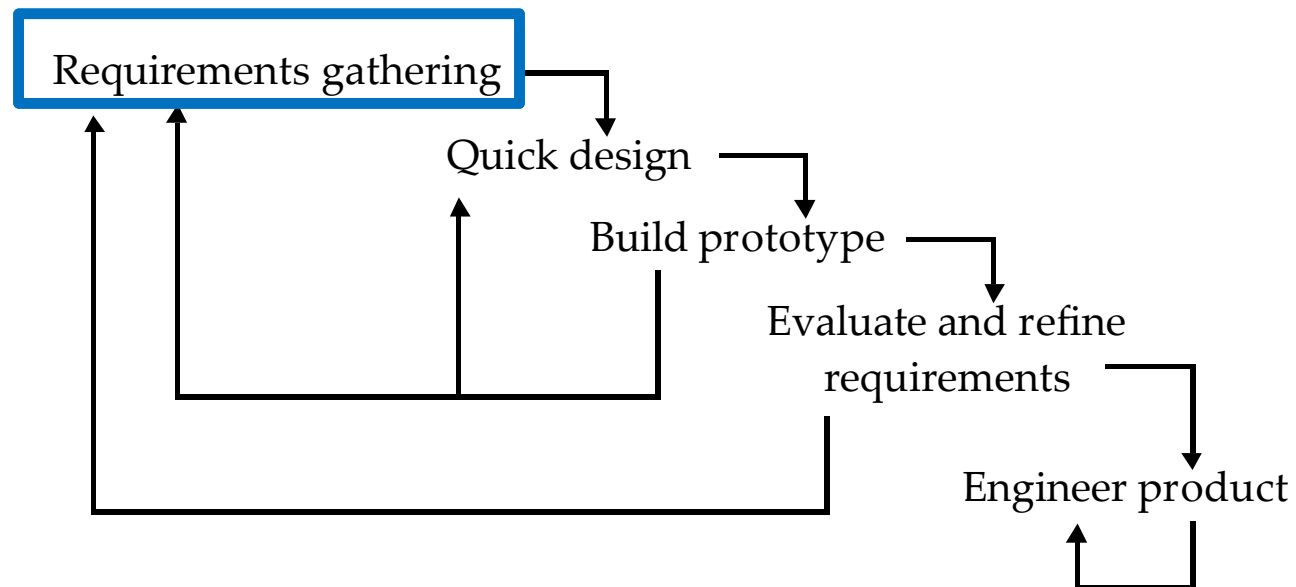
# Interaction Design



Gathering and Analysing Data

# Interaction Design

- **Iterative** user centered design and development





# This Lecture

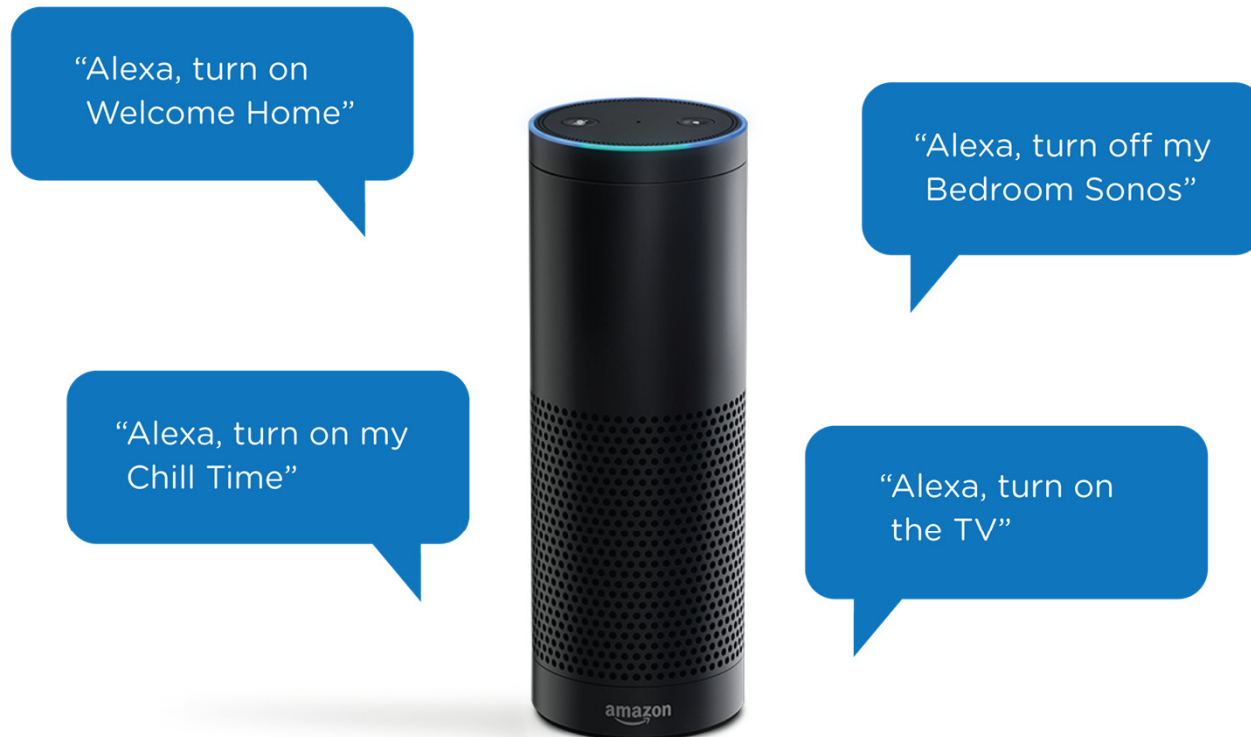
- **Data gathering**

- Observation
- Interviews
- Focus groups
- Card sorting
- Questionnaires
- Studying documentation
- Scenarios / Use cases
- Researching similar products
- Web analytics

- **Data analysis & interpretation**

- Quantitative data analysis
- Qualitative data analysis

# Case Study: Amazon Alexa



These four slides are based on: **Amazon Alexa talk by Dr Craig Saunders** (31 January 2018)

# Alexa: Single vs. Multiple Answers

- Should Alexa give one single answer to a user's question or multiple answers?
- How does Amazon make such a decision?

# Alexa: Asking Questions

- Should Alexa ask you a question without you calling her to action by saying 'Hey Alexa'?
- How does Amazon know that?

# Alexa: Sensitive Topics

- Does Alexa know which topics are sensitive and should not be talked about / discussed?
- How does Amazon know that?

# Gathering Data







# Gathering Data

- To collect **sufficient, relevant, and appropriate** data
  - to produce a stable set of requirements
- To collect data about users, their tasks, the context of use, and the rationale



# Data Collection Techniques

- Observation
- Interviews
- Focus groups
- Card sorting
- Questionnaires
- Studying documentation
- Scenarios / Use cases
- Researching similar products
- Web analytics

# Observation (1)

- Observation **in the field** (ethnography)
  - ⊕: Realistic settings/ activities
  - ⊗: Difficult to set up
  - ⊗: Complex observation may intrude upon the user:
    - May effect their performance
    - May skew the data
  - ⊗: Problems of privacy/ reliability
- Observation **in the lab**
  - ⊕: Less intrusion
  - ⊗: But not realistic setting/ activities

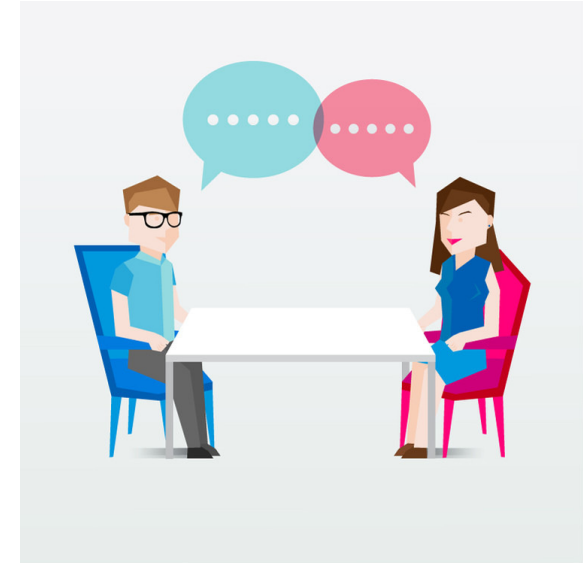


# Observation (2)

- **Direct** observation: Concurrent protocol
  - The users say what they are doing and why, whilst they are performing the task
    - ⊕: good for understanding the nature and context of the task
    - ⊗: too much data, time consuming
- **Indirect** observation: Retrospective protocol
  - The users explain what they did and why, after they have finished the event
    - e.g. observing the users through records of their activity such as logs

# Interviews (1)

- An interviewer asks an interviewee a set of questions
- May be structured in different ways:
  - **Unstructured interviews** (open questions)
    - ⊕: rich data, ⊗: too much data, off target, long
  - **Structured interviews** (closed questions)
    - ⊕: easy data collection, ⊗: rigid
  - **Semi-structured interviews** (open & closed questions)
    - ⊕: rich targeted data



# Interviews (2)

- Attempt to identify user's / interviewee's subjective opinions
  - ⊕: forum for talking to people
  - ⊕: good for exploring issues
- **Avoid:**
  - Long questions
  - Compound sentences - split them into two
  - Jargon and language that the interviewee may not understand
  - Leading questions that make assumptions e.g.. why do you like ...?
  - Unconscious biases e.g.. gender stereotypes

# Interviews (3): Enriching the Process

- Props - devices for prompting interviewee, e.g. use a prototype, scenario



# Focus Groups

- Interviewing people in groups
  - Used to identify conflicts in terminology or expectations from different groups
  - **Important**: select participants to represent well the target users



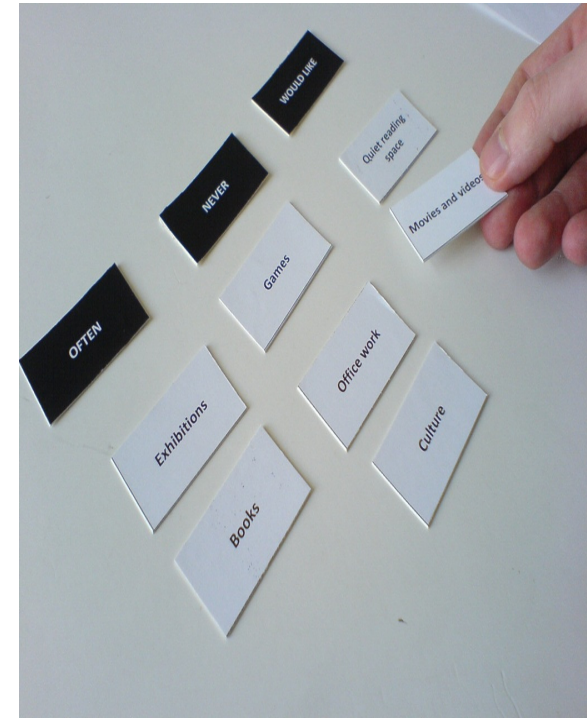


# Card Sorting (1)

- Applied to a wide variety of activities involving the **grouping** and/or **naming** of objects or concepts
- The sorting process is performed by potential users of an interactive solution
- It provides:
  - Terminology (what people call things)
  - Relationships (proximity, similarity)
  - Categories (groups and their names)

# Card Sorting (2)

- Participant is given a set of paper cards
    - Each card has some aspect of the work or environment on it
  - Participant is asked to sort the cards into groups
    - Typically groups of similar items
- ⊕: Can get an idea of participant's perceptions of the nature of the items and activities



# Card Sorting (3)

- **Open Card Sorting:** Participants are given cards showing content with no pre-established groupings
  - They are asked to sort cards into groups that they feel are appropriate and then describe each group
- **Closed Card Sorting:** Participants are given cards showing content with an established initial set of primary groups
  - Participants are asked to place cards into these pre-established primary groups
- **Hybrid Card Sorting:** Some combination of the two

# Case Study: Open Card Sorting

- Your team is responsible for the information architecture of the self-checkout counters of a large supermarket chain
  - There have been some complaints that customers are spending a long time at these counters and are frustrated by how the categories are organized



<https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/card-sorting>

# Case Study: Open Card Sorting

- You are asked to consider how you might organize these yourself
  - Use the items listed below and group them into **four groups** based on **similarity**
  - Give a name/title for each group
    - apples
    - oranges
    - strawberries
    - bananas
    - peaches
    - plums
    - tomatoes
    - pears
    - grapes
    - cherries





# Open Card Sorting: **Solution**

# Card Sorting: Example

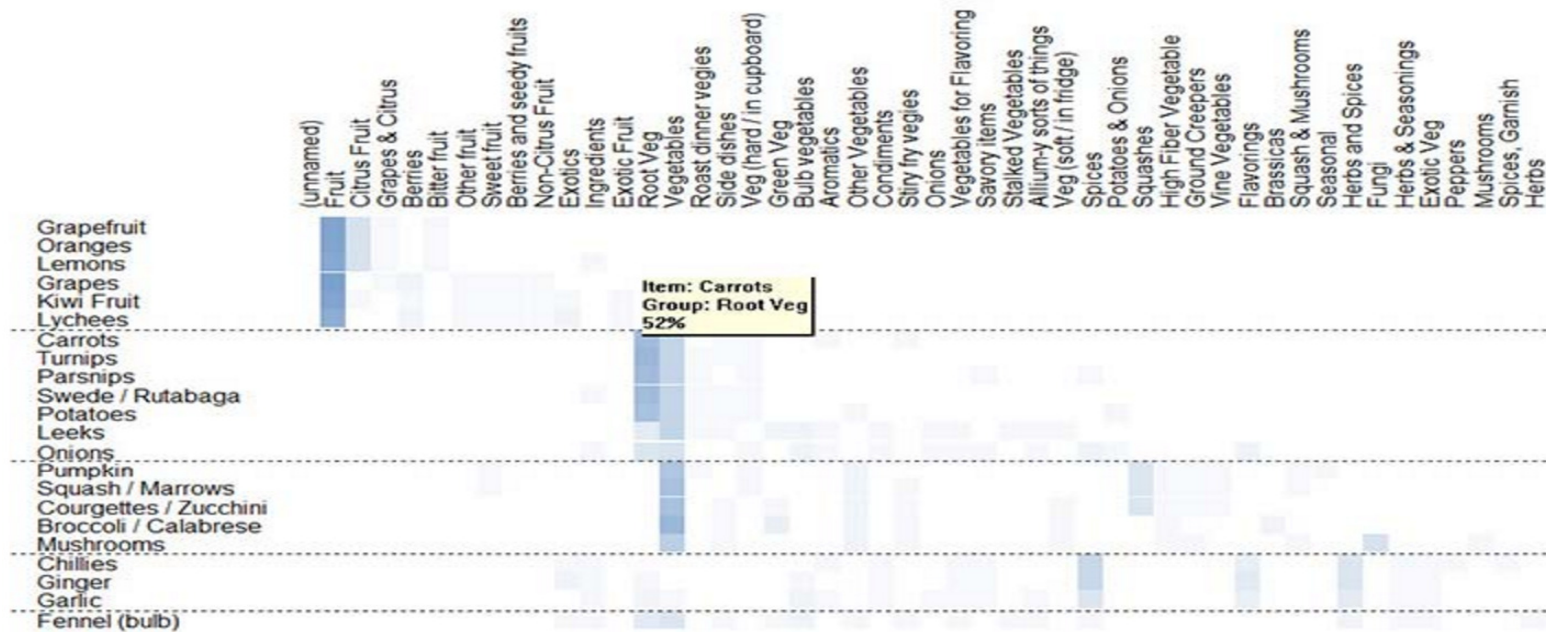


Figure 22.5: Items by groups chart of fruit & vegetables sample with 26 participants (SynCaps V2)



# Open Card Sorting

## A step-by-step guide

1. Get a pack of (index) cards
2. Use 15~40 cards (size depends on the project: e.g. some projects used 78 cards!)
3. List each content item on a separate card
  - Avoid using terms that can be mistaken as a category
  - Avoid using phrasing that will tend to imply a group
4. Number the cards on the back
5. Scatter cards on a desk in a random order
6. Ask the user to sit at the desk and sort the cards into piles of related items
7. After the cards are sorted, ask the user to give a name to each pile
8. After the interview, turn the piles over and use the numbers on the back of the cards to record the groupings
9. Record the name of the group given by the user
10. Repeat this process with a number of users



# Questionnaires (1)



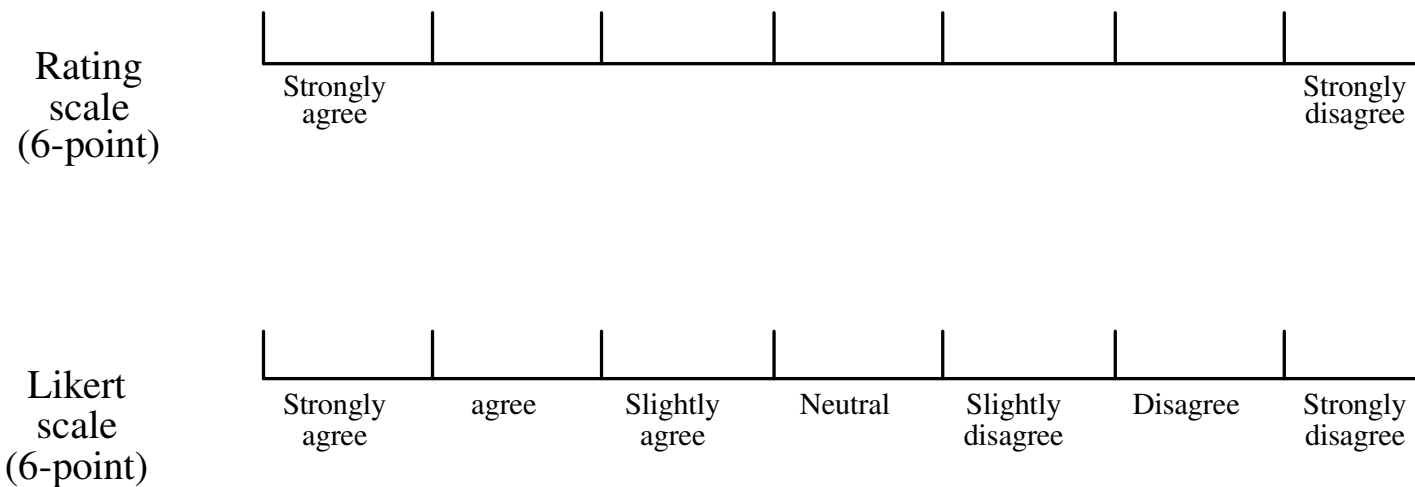
- A series of questions designed to elicit **specific** information
- Good for getting answers to specific questions from a large, dispersed group of people
- Can provide quantitative or qualitative data
- Often used **in conjunction** with other techniques

# Questionnaires (2)

- **Open questions** – the respondent is free to write their answers in any way they see fit
- **Closed questions** – the respondent selects an answer from a set of presented possibilities
  - Several kinds of scale from which participants can choose their response including:
    - **Simple checklist** – simple responses such as 'yes', 'no', or 'don't know' are provided
    - **Multi-point rating scale** – a number of points on a scale are provided which give a range of responses
    - **Ranked order** – respondent specifies their preference for items in a list

# Questionnaires (3)

- Questionnaires: Multi-point ranking
  - E.g. a variation of the Likert scale



# Questionnaires: Example

- Questionnaire to evaluate two different phone designs

*For each pair of adjectives, place a cross at the point between them that reflects the extent to which you believe the adjectives describe the phone design. Please place only one cross between the marks on each line.*

*Annoying*

--	--	--	--	--	--

*Pleasing*

*Easy to use*

--	--	--	--	--	--

*Difficult to use*

*Value-for-money*

--	--	--	--	--	--

*Expensive*

*Attractive*

--	--	--	--	--	--

*Unattractive*

*Secure*

--	--	--	--	--	--

*Not secure*

*Helpful*

--	--	--	--	--	--

*Unhelpful*

*Hi-tech*

--	--	--	--	--	--

*Lo-tech*

*Robust*

--	--	--	--	--	--

*Fragile*

*Inefficient*

--	--	--	--	--	--

*Efficient*

*Modern*

--	--	--	--	--	--

*Dated*

# Questionnaires: Things to Remember (1)

- The impact of a question can be influenced by **question order**
- You may need **different versions** of the questionnaire for different populations
- Provide **clear instructions** on how to complete the questionnaire
- **Avoid** very **long questionnaires**
- Decide on whether **phrases will all be positive**, all negative or mixed

# Questionnaires: Things to Remember (2)

- Make sure **purpose of study** is clear
- Promise **anonymity**
- Ensure questionnaire is **well designed**
- Offer a **short version** for those who do not have time to complete a long questionnaire
- Provide an **incentive**
- 40% response rate is good, 20% is often acceptable

# Questionnaires: Example (2)

## QUESTIONNAIRE FOR PARENTS OF AFFECTED HEARING IMPAIRMENT CHILD

Name \_\_\_\_\_ Education \_\_\_\_\_

Profession \_\_\_\_\_ Annual Income:  50,000p.a  1lakhp.a  1lakhp.a  2lakhp.a

No. of children \_\_\_\_\_ Affected Child \_\_\_\_\_

Delivery of child:  Normal  Caesarean  Forceps Any health Problem during Pregnancy:  Yes  No

Family History \_\_\_\_\_

How did they identified about the deafness of their child \_\_\_\_\_

at what age \_\_\_\_\_ and what you did \_\_\_\_\_

Degree of deafness:  Normal  Mild  Moderate  Severe  Profound

What hearing aid you got:  Pocket  Inside the canal  Behind the ear  In the ear  Completely in ear

Did you make full use of facilities provided by the government:  Yes  No

How much time did the child took adapting the hearing aid:  2 week  1 month  more than 1 month

How did you approach Speech therapist \_\_\_\_\_

How many times per week did you take your child to speech therapist \_\_\_\_\_

What was your teaching methodology adopted :  
 Formal  Informal (nature or learn while fun)

Which language did you choose initially :  
 Mother tongue  English  Hindi

Types of study material : (Alphabet dictionary, animal, fruits, vegetable, bird, transport, etc.)  
 Book  Chart  Word and image  Image  Word  Other \_\_\_\_\_

Kind of techniques used :  
 Drawing  Coloring  Cutting pasting  Clay Modeling  Other \_\_\_\_\_

Kind of stories book :  
 Pictorial (Stories in sequence order)  Pictorial and word  Word  Other \_\_\_\_\_

Kinds of toys :  
 Puzzles  Building blocks  Video games  Ludo  Scrabble  Mind games  Other \_\_\_\_\_

Kind of Audio cassettes :  
 Rhyme  Narrating story  Bollywood songs  Other \_\_\_\_\_

Any specialized training in speech therapist taken by you and from where ? \_\_\_\_\_

How much time did you spend with the child everyday? \_\_\_\_\_

How you got involved with the child to teach him language? \_\_\_\_\_

How was the progress of child in vocabulary development, clarity of speech, pronunciation?  
 Poor  Average  Good  Very Good  Excellent

Did you send their child to normal school or special school, at what age? \_\_\_\_\_

If normal schools, what difficulty did the child face during the class and how did the child overcome? \_\_\_\_\_

How was the teacher involvement with the affected child in schools?  
 Poor  Average  Good  Very Good  Excellent

# Studying Documentation



- Procedures and rules are often written down in manuals
  - Good source of data about the steps involved in an activity
    - and regulations governing a task
  - Good for understanding legislation, and getting background information
  - No stakeholder time
    - a limiting factor on the other techniques
  - Not to be used in isolation
- Overall:
  - ⊕: availability, accessibility
  - ⊗: idealized reports, outdated



# Researching Similar Products

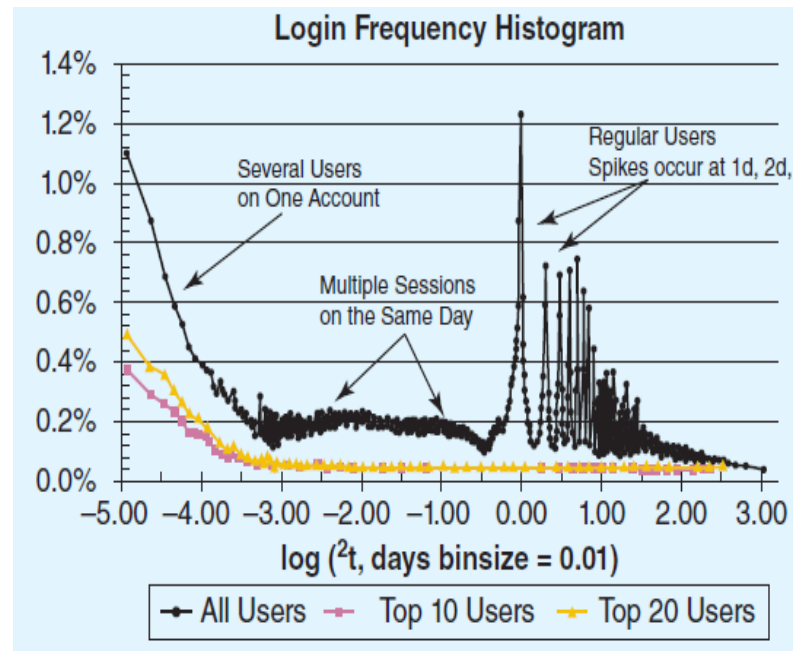
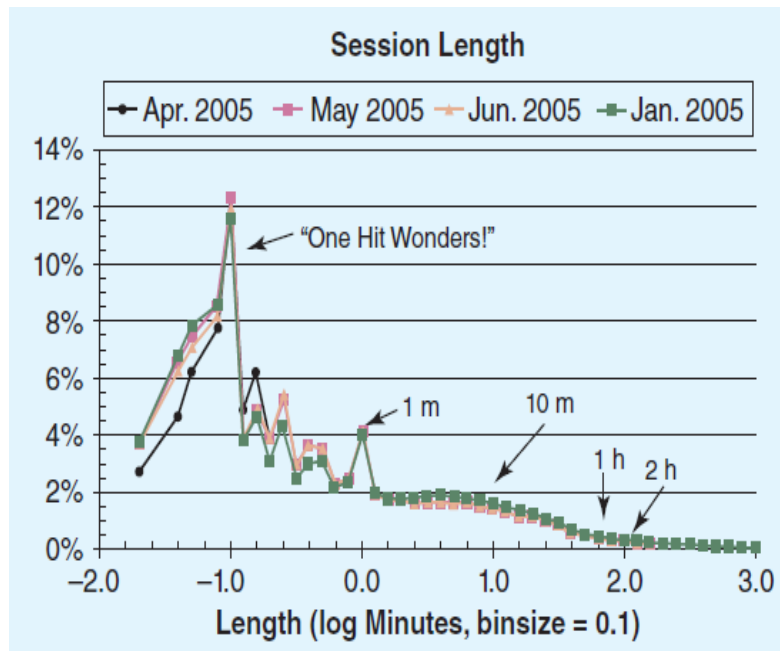
- ⊕ Helps to prompt requirements
- ⊕ Helps to generate alternative designs
- ⊗ May inhibit creativity



# Web Analytics

- A system of tools and techniques for optimizing web usage by:
  - Measuring
  - Collecting
  - Analysing
  - Reporting web data
- Typically focus on the number of web visitors and page views

# Web Analytics: Example



session length data of four different months from Teachers' Domain (NSDL)

Source: Khoo, M., Pagano, J., Washington, A. L., Recker, M., Palmer, B., and Donahue, R. A. (2008) Using web metrics to analyze digital libraries. *Proceedings of Joint Conference on Digital Libraries*, Pittsburgh, June 16–20. ©2008 Association for Computing Machinery, Inc. Reprinted by permission.

# Choosing Between Techniques

- The choice of technique depends on:
  - Amount of **time**, level of **detail** and **risk** associated with the findings
  - **Knowledge** the analyst requires
- The choice is also affected by the **kind of task** studied:
  - Sequential steps or overlapping series of subtasks?
  - High or low, complex or simple information?
  - Task for a layman or a skilled practitioner?

# Basic Guidelines (1)

- Focus on identifying the stakeholders' needs
- Involve all the stakeholder groups
- Involve more than one representative from each stakeholder group
- Triangulation
  - Use a combination of data gathering techniques
  - Collect more than one type of data

## Basic Guidelines (2)

- Consider carefully how to record the data
- You will need to compromise on the data you collect and the analysis to be done
- Support the process with **prototypes** and **task descriptions**
- Balance functional and non-functional requirements

# Data Analysis & Interpretation



# Data Analysis & Interpretation

- Helps you interpret and present your findings in appropriate ways
- Qualitative and quantitative data analysis
  - To analyse data gathered from:
    - Questionnaires
    - Interviews
    - Observation studies
    - ...



# Quantitative vs. Qualitative Data

- Quantitative analysis
  - numerical methods to ascertain size, magnitude, amount
- Qualitative analysis
  - expresses the nature of elements
  - represented as themes, patterns, stories
  - difficult to measure sensibly as numbers
    - e.g. count number of words to measure dissatisfaction

# Initial Processing

	Usual raw data	Example qualitative data	Example quantitative data	Initial processing steps
Interviews	Audio recordings. Interviewer notes. Video recordings	Responses to open questions. Video pictures. Respondent's opinions	Age, job role, years of experience. Responses to closed questions	Transcription of recordings. Expansion of notes
Questionnaires	Written responses. Online database	Responses to open questions. Responses in 'further comments' fields. Respondent's opinions	Age, job role, years of experience. Responses to closed questions	Clean up data. Filter into different data sets
Observation	Observer's notes. Photographs. Audio and video recordings. Data logs. Think-aloud	Records of behavior. Description of a task as it is undertaken. Copies of informal procedures	Demographics of participants. Time spent on a task. The number of people involved in an activity	Expansion of notes. Transcription of recordings. Synchronization between data recordings

# Initial Processing: Example

- Questionnaire to evaluate two different phone designs

*For each pair of adjectives, place a cross at the point between them that reflects the extent to which you believe the adjectives describe the phone design. Please place only one cross between the marks on each line.*

*Annoying*

--	--	--	--	--	--

*Pleasing*

*Easy to use*

--	--	--	--	--	--

*Difficult to use*

*Value-for-money*

--	--	--	--	--	--

*Expensive*

*Attractive*

--	--	--	--	--	--

*Unattractive*

*Secure*

--	--	--	--	--	--

*Not secure*

*Helpful*

--	--	--	--	--	--

*Unhelpful*

*Hi-tech*

--	--	--	--	--	--

*Lo-tech*

*Robust*

--	--	--	--	--	--

*Fragile*

*Inefficient*

--	--	--	--	--	--

*Efficient*

*Modern*

--	--	--	--	--	--

*Dated*

# Initial Processing: Example

- Questionnaire to evaluate two different phone designs
  - 100 respondents for the evaluation of phone 1

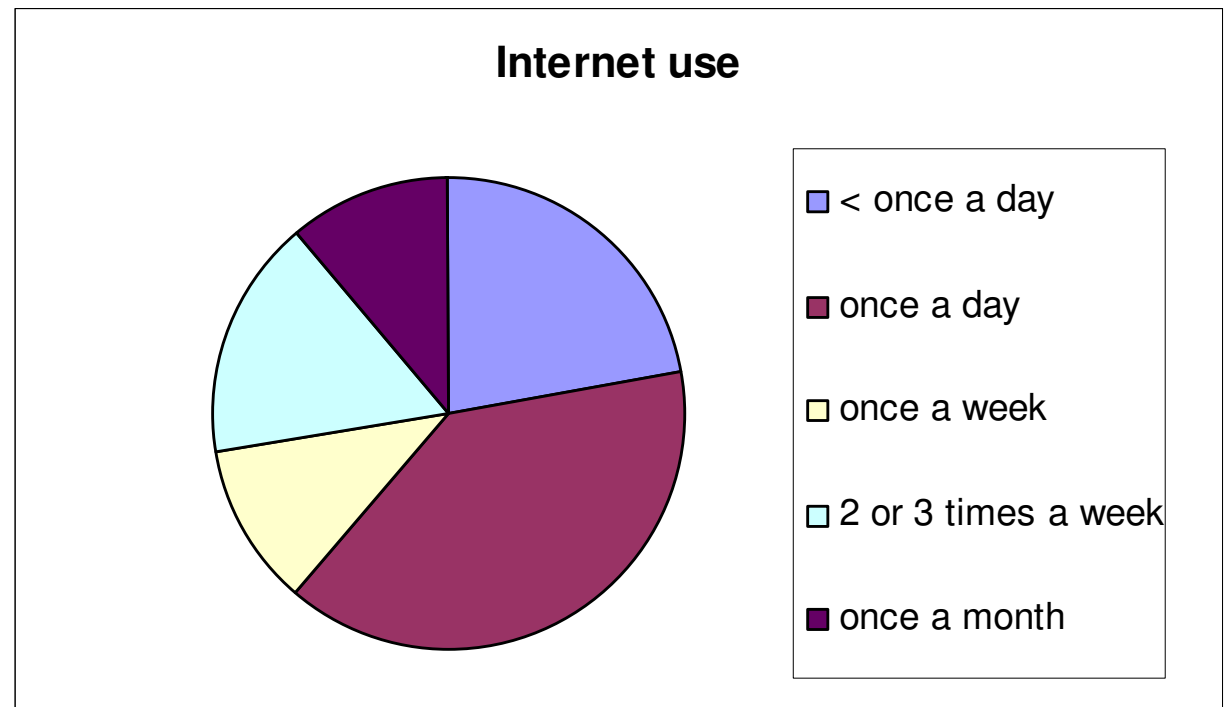
	1	2	3	4	5	
Annoying	35	20	18	15	12	Pleasing
Easy to use	20	28	21	13	18	Difficult to use
Value-for-money	15	30	22	27	6	Expensive
Attractive	37	22	32	6	3	Unattractive
Secure	52	29	12	4	3	Not secure
Helpful	33	21	32	12	2	Unhelpful
Hi-tech	12	24	36	12	16	Lo-tech
Robust	44	13	15	16	12	Fragile
Inefficient	28	23	25	12	12	Efficient
Modern	35	27	20	11	7	Dated

# Simple Quantitative Analysis

- Averages
  - Mean: add up values and divide by number of data points
  - Median: **middle value** of data when ranked
  - Mode: figure that appears **most often** in the data
- Percentages
- Graphical representations give overview of data
- **Important**: Be careful not to mislead with numbers!

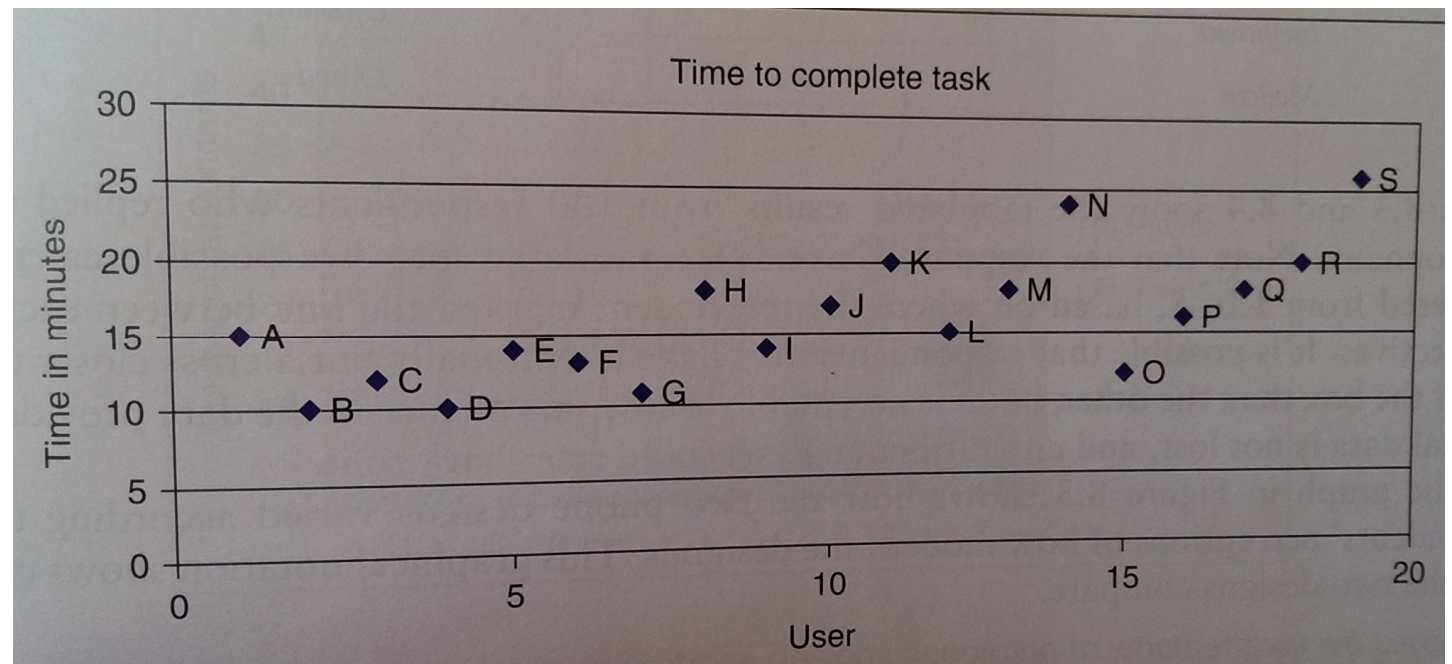
# Quantitative Analysis: Example (1)

- Results:
  - Experience of users



# Quantitative Analysis: Example (2)

- Results:
  - Time to complete the task



# Quantitative Analysis: Example (3)

- An evaluation study of a document sharing application
  - The experience of the users and the number of errors made while trying to complete the controlled task

User	Internet use					Number of errors made
	More than once a day	Once a day	Once a week	Two or three times a week	Once a month	
1		1				4
2	1					2
3			1			1
4	1					0
5				1		2
6		1				3
7	1					2
8		1				0
9					1	3
10	1					2
11				1		1
12			1			2
13		1				4
14		1				2
15						1
16				1		1
17		1			1	0
18		1				0
Totals	4	7	2	3	2	30
					Mean	1.67 (to 2 decimal places)



# Quantitative Analysis : Card Sorting

- Need to find commonalities of how people grouped the cards
  - Which items were grouped together most often?
- Small projects
  - Just looking at the piles for patterns could be sufficient
- Larger projects
  - Will require some form of analysis, ranging from simple tabulation through to [cluster analysis](#)

# Card Sorting: Example of Analysis

## Fruits example

- Results from 1 participant

A	B	C	D	E	F	G	H	I	J	K
	apples	oranges	strawberries	bananas	peaches	plums	tomatoes	pears	grapes	cherries
apples	--	0	1	1	0	1	0	1	1	1
oranges	0	--	1	1	0	1	0	1	1	1
strawberries	1	1	--	1	1	0	1	1	0	0
bananas	1	1	1	--	1	1	1	0	1	1
peaches	0	0	1	1	--	1	0	1	1	1
plums	1	1	0	1	1	--	1	1	0	0
tomatoes	0	0	1	1	0	1	--	1	1	1
pears	1	1	1	0	1	1	1	--	1	1
grapes	1	1	0	1	1	0	1	1	--	0
cherries	1	1	0	1	1	0	1	1	0	--

The entries in this spreadsheet indicate whether this participant put each pair of fruits into the same or different groups. A "0" means the same group and a "1" means different groups.

# Card Sorting: Example of Analysis

## Fruits example

- Distance matrix: results from 20 participants

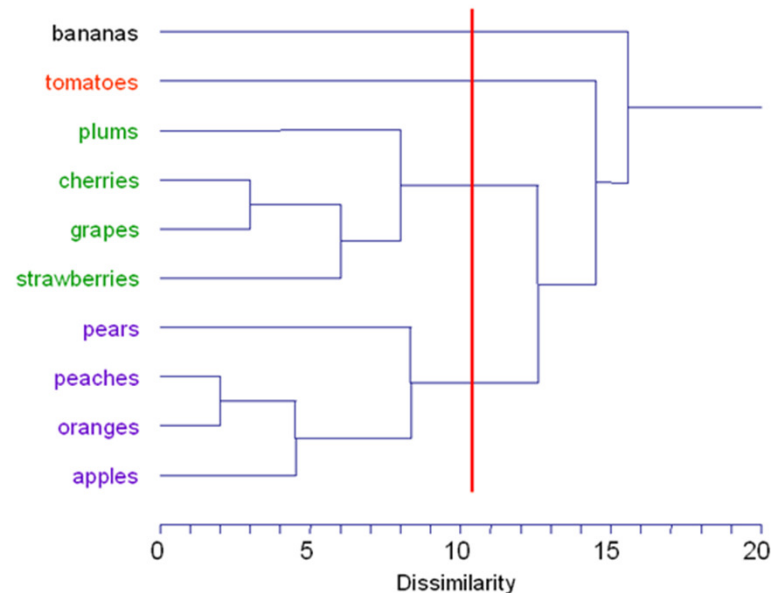
A	B	C	D	E	F	G	H	I	J	K
	apples	oranges	strawberries	bananas	peaches	plums	tomatoes	pears	grapes	cherries
apples	--	5	11	16	4	10	12	8	11	10
oranges	5	--	17	14	2	12	15	11	12	14
strawberries	11	17	--	17	16	8	18	15	4	8
bananas	16	14	17	--	17	15	20	11	14	16
peaches	4	2	16	17	--	9	11	6	15	13
plums	10	12	8	15	9	--	12	10	9	7
tomatoes	12	15	18	20	11	12	--	16	18	14
pears	8	11	15	11	6	10	16	--	12	14
grapes	11	12	4	14	15	9	18	12	--	3
cherries	10	14	8	16	13	7	14	14	3	--

This spreadsheet is a "distance matrix" created by summing all of the individual spreadsheets from the participants in the card-sorting study. In this example, we had 20 participants. The "oranges-peaches" pair has a value of 2, for example, indicating that only 2 participants put them in different groups. The "tomatoes-bananas" pair, on the other hand, has a value of 20, indicating that all 20 participants put them in different groups.

# Card Sorting: Example of Analysis

## Fruits example

- Results of **hierarchical cluster analysis** of the sample data in the form of dendrogram
  - a tree diagram showing taxonomic relationships



Further reading:

<http://www.econ.upf.edu/~michael/stanford/maeb7.pdf>

# Card Sorting : Cluster Analysis

- Measure: Similarity rating
  - Every time two cards are in the same pile you assign them 1 point
  - **Similarity rating** = add up all of the times that two cards appear together and divide by the number of groups

# Cluster Analysis: Example

- Calculate the similarity rating

Card-sorting results		User A	User B	User C	User D					
		1,5	1,4,5	1,5	1					
		2,3	2,3	2,3	2,3,4,5					
		4		4						
Score										
cards 1 and 2		0	+	0	+	0	+	0	=	0
cards 1 and 3		0	+	0	+	0	+	0	=	0
cards 1 and 4		0	+	1	+	0	+	0	=	1
cards 1 and 5		1	+	1	+	1	+	0	=	3
Similarity rating										
cards 1 and 2		0/4 = 0								
cards 1 and 3		0/4 = 0								
cards 1 and 4		1/4 = .25								
cards 1 and 5		3/4 = .75								

# Qualitative Data Analysis

- **Recurring** patterns or themes
  - Emergent from data
  - Dependent on observation framework if used
- Categorizing data
  - Categorization scheme may be **emergent** or **pre-specified**
- Looking for critical incidents
  - Helps to focus in on key events



# Qualitative Data Analysis

- Theoretical Frameworks
  - Basing data analysis around theoretical frameworks provides further insight
  - Such frameworks are:
    - Grounded Theory
    - Distributed Cognition
    - Activity Theory
    - Thematic Analysis



# Tools to Support Data Analysis

- Spreadsheet – simple to use, basic graphs
- Statistical packages
  - e.g. SPSS
- Qualitative data analysis tools
  - Categorization and theme-based analysis
  - Quantitative analysis of rich text-based or multimedia data
  - Example software for qualitative data analysis:
    - Nvivo
    - Atlas.ti
    - CAQDAS Networking Project, based at the University of Surrey  
<http://www.surrey.ac.uk/sociology/research/researchcentres/caqdas/>

# Presenting the Findings

- The best way to present your findings depends on
  - the audience
  - the purpose
  - the data gathering and analysis undertaken
- Use graphical representations as needed
- **Important:** Only make **claims that your data can support !**

# Summary

- Data collection is critical to successful design
- The data analysis that can be done depends on the data gathering that was done
- Choose your data collection techniques wisely
  - Kinds of data collected
  - Time to collect and interpret data
  - Interference with activity being studied
- Presentation of the findings should not overstate the evidence

# Study Material & Reading

- BOOK: Preece, J., Rogers, Y. and Sharp, H. **Interaction Design.**
  - Chapter: Data Gathering
  - Chapter: Data Analysis, Interpretation, and Presentation
  - Chapter: Establishing Requirements