TYPES OF DATA
QUALITATIVE ANALYSIS TECHNIQUE

CPSC 544 FUNDAMENTALS IN DESIGNING INTERACTIVE COMPUTATION TECHNOLOGY FOR PEOPLE (HUMAN COMPUTER INTERACTION)

WEEK 4 – CLASS 6

© Joanna McGrenere
Includes slides from Leila Aflatoony, Karon MacLean and Jessica Dawson
TODAY

• Admin [5min]
• Types of data [5min]
• Qualitative analysis techniques [25min]
• In class activity [20min]
  • coding
• Discussion of readings [20min]
• Project [5min]
LEARNING GOALS

• explain types of data that HCI methods provide
• explain what it means to triangulate in data gathering and analysis
• understand theme-based approaches to analysis
• understand thematic analysis as one of the most common forms of analysis and describe the steps to doing thematic analysis
• understand why reliability and validity are important in qualitative research
TYPES OF DATA THAT HCI METHODS PROVIDE

spectrum of data...

qualitative:

• users describe/report X, to extent they are aware
• you observe X, that users may not be fully aware of

where X can be: behaviours, processes, usability challenges…

quantitative:

• measure task performance with existing tools / methods:
  • e.g., speed, errors, dead-ends, learning curves for novices …
• numerical data from user-reported answers: e.g., # of emails/day
• counting observed occurrences: e.g. # of times looked at instructions
1. controlled observations – of time to complete task?
2. unstructured observations – of observed steps to complete a task?
3. unstructured interviews – of user telling stories?
4. questionnaire – self report numbers of times switch to social media/day?
TRIANGULATION

a strategy to enhance validity:
use the *multiple perspectives* available from complementary sources

use multiple:
- data sources
  - people, places, times
- data collection methods
- researchers/evaluators

image credit: Sandra Mathison, UBC EPSE 595
ANALYZING & INTERPRETING DATA

qualitative data – interpreted to tell a “story”
  • categories, themes, patterns, etc.

quantitative data – presented as values, tables, charts and graphs
  • counts (e.g., summary of total # of errors)
  • simple statistical analysis (e.g., averages)
  • advanced statistical analysis (e.g., linear regression)

… more on quantitative coming later this term.

your choice of analysis method will depend on what you’re using it for
  • remember you can often transform from qualitative to quantitative
METHODS FOR QUALITATIVE ANALYSIS

From simpler (less effort) to more advanced (more effort):

• identify critical incidents
• simple categorization
• categorization - themes, patterns

In all cases your aim is to interpret the data in ways that encapsulate and document your understanding.

Level of effort depends on your goal.

Many methods often used in combination.
QUALITATIVE ANALYSIS METHODS
FINDING CRITICAL INCIDENTS

identify and focus on the most significant incidents

- efficient when you have lots of data
- incidents can be either desirable or undesirable

NOT about summarizing all incidents that occur

→ more like finding tiny gold nuggets in buckets of sand
→ appropriate for usability studies, where qualitative is not primary analysis

incidents

non (or less) relevant data
QUALITATIVE ANALYSIS METHODS

CATEGORIZING DATA

typically used on transcripts (observations, interviews, etc.)

- at non-granular level of detail to find stories or themes
- at fine-grain level of detail focusing on words, sentences, gestures, etc. (e.g., discourse analysis)

first data are ‘coded’ according to a scheme of categories

- can be predetermined (deductive), or arise from the data (inductive)
  - at a high level: affinity diagram
EXAMPLE OF CODING A TRANSCRIPT MID- TO LOW-LEVEL OF DETAIL

Very informative, but time consuming!
THEMATIC ANALYSIS

• Thematic analysis is one of the most common forms of analysis in qualitative research.

• It is a method for identifying, analysing and reporting patterns (themes) within data. It minimally organizes and describes your data set in (rich) detail.

QUALITATIVE ANALYSIS METHODS

PATTERNS AND THEMES

can be revealed in many ways:

• through the process of conducting the study
• use of tools and techniques (e.g., affinity diagrams)
  • will talk about it more on Thursday

can support many types of user study goals

• e.g., understanding behaviour, culture, places or situations where events occur, breakdowns, user characteristics, etc.

very flexible and widely used

• can be reported as findings or inform more analysis
INDUCTIVE VS. THEORETICAL THEMATIC ANALYSIS

• Inductive or ‘bottom up’ way
  • data-driven; process of coding the data without trying to fit it into a preexisting coding frame, or the researcher’s analytic preconceptions

• Theoretical or deductive or ‘top down’ way
  • analyst driven; driven by the researcher’s theoretical or analytic interest in the area
DOING THEMATIC ANALYSIS: A STEP-BY-STEP GUIDE

1. Familiarizing yourself with your data
2. Generating initial codes
3. Searching for themes
4. Reviewing themes
5. Defining and naming themes
6. Producing the report
1. FAMILIARIZING YOURSELF WITH YOUR DATA

Transcribing data, reading and re-reading the data, noting down initial ideas.

Note:

• writing is an integral part of analysis
• analysis is not a linear process
• read through the entire data set before you begin your coding
2. GENERATING INITIAL CODES

Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code.

- your coded data differ from the units of analysis (your themes), which are (often) broader.
  - Note: that your reading uses “themes” for narrow, and “codes” for broad categories -> less standard terminology
- coding will, to some extent, depend on whether the themes are more ‘data-driven’ or ‘theory-driven’

CODING MANUALLY

Code your data by writing notes on the texts you are analysing, by using highlighters or coloured pens to indicate potential patterns, or by using ‘post-it’ notes to identify segments of data.

Key advice for this phase is:

(a) code for as many potential themes/patterns as possible
(b) code extracts of data inclusively / i.e., keep a little of the surrounding data if relevant
CODING WITH COMPUTER SOFTWARE

If using computer software, you code by tagging and naming selections of text within each data item.

- NVivo: http://www.qsrinternational.com/nvivo-product
- ATLAS.ti: http://atlasti.com
- Saturate App: http://www.saturateapp.com
3. SEARCHING FOR THEMES

Collating codes into potential themes, gathering all data relevant to each potential theme.

• Analyse codes and consider how different codes may combine to form an overarching theme.

• It may be helpful at this phase to use visual representations to help you sort the different codes into themes. You might use tables, or mind-maps, or write the name each code.
4. REVIEWING THEMES

Checking if the themes work in relation to the coded extracts and the entire data set, generating a thematic ‘map’ of the analysis.

Goal:

- to ascertain whether the themes ‘work’ in relation to the data set.
- to code any additional data within themes that has been missed in earlier coding stages.

Note:

Data within themes should cohere together meaningfully, while there should be clear and identifiable distinctions between themes.
5. DEFINING AND NAMING THEMES:

Identify the ‘essence’ of what each theme is about

- Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.
- Identify whether or not a theme contains any sub-themes
- Give names to the themes

6. PRODUCING THE REPORT

The final opportunity for analysis.

• Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to the research question and literature, producing a scholarly report of the analysis.

• It is important that the analysis (the write-up of it, including data extracts) provides a concise, coherent, logical, non-repetitive and interesting account of the story the data tell/within and across themes.

Reliability and validity are fundamental concerns of the qualitative researchers

- Transparency of technique
  - Carefully documenting all their steps so that they can be checked by another researcher
- Reliability checker
  - Organizing an independent assessment of transcripts by additional skilled qualitative researchers and comparing agreement between the raters.
  - Can be done statistically (called inter-rater reliability) or qualitatively by discussing disagreements
  - Note: different schools of thought on need for inter-rater reliability
COMMON PITFALLS TO THEMATIC ANALYSIS

1. Failure to analyse: TA is not just a collection of extracts strung together with little or no analytic narrative.
2. Using data collection (e.g., interview) questions as the themes that are reported.
3. Weak or unconvincing analysis, e.g., too much overlap between themes
4. Mismatch between the data and the analytic claims
5. Mismatch between theory and analytic claims
IN CLASS ACTIVITY [20 MIN]

• First steps in coding a transcript
Hoarding and Minimalism: Tendencies in Digital Data Preservation

Francesco Vitale, Izabelle Janzen, Joanna McGrenere
University of British Columbia, Vancouver, Canada
fvitale, bjanzen, joanna@cs.ubc.ca

ABSTRACT
Digital data, from texts to files and mobile applications, has become a pervasive component of our society. With seemingly unlimited storage in the cloud at their disposal, how do people approach data preservation, deciding what to keep and discard? We interviewed 23 participants with diverse backgrounds, asking them about their perceived digital data: what “stuff” they kept through the years, why, how they used it, and what they considered important. In an iterative analysis process, we uncovered a spectrum of tendencies that drive preservation strategies, with two extremes: hoarding (where participants accumulated large amounts of data, even if considered of little value) and minimalism (where they kept as little as possible, regularly cleaning their data). We contrast and compare the two extremes of the spectrum, characterize their nuanced nature, and discuss how our categorization compares to previously reported behaviors such as filing and piling, email cleaners and keepers. We conclude with broad implications for shaping technology.

ACM Classification Keywords
H.3.m Information Storage and Retrieval: Miscellaneous

Author Keywords

side. This explains why they are an increasingly popular choice to store digital data for everyday users [44]. Storage is either cheap or outright free. Google Photos, for example, offers unlimited space for pictures (although at reduced quality).

This is the “seductive” digital landscape that Marshall [25] predicted a decade ago when studying long-term preservation of digital items. At the time, a similar change was taking place: hard drive storage was becoming cheaper, giving users the option to store nearly “everything” [26]. The pervasiveness of the cloud is once again reinforcing this possibility. Now that we are living in this seductive landscape, how are data preservation practices changing? It is critical to understand how users are experiencing this new world, as we are just in its foothills. As storage gets cheaper and digital data more of a commodity, how do users deal with this new environment?

We are interested in the act of preserving data, by which we mean deciding what data to keep and discard. As Whittaker [47] points out: little is known about “when and why people keep or delete different types of information.” Therefore, we focused on a main, broad research question: how do people approach digital data preservation in the cloud age? How do they decide what to keep and discard?
DISCUSSION ON DATA ANALYSIS READINGS [20 MIN]

Get into group of 3-4 answering the following questions:

- What surprised you? or
- What you disagreed with?
- Others?
HIGHERY RECOMMENDED COURSE

If you are going to use qualitative methods in any significant way in your research...

UBC EPSE 595: Qualitative Research Methods
http://ecps.educ.ubc.ca/epse-595-qualitative-research-methods/
PROJECT QUESTIONS [5 MIN]

• We put brief comments on project interim milestone on Canvas
• Stay after class to discuss any concerns
• Remember:
  • need consent forms for interviews and questionnaires
  • need to follow HCI Course Ethics
  • need to get consent and instruments (questions) quickly approved by course staff before using
ON DECK...

Next class (Thursday) …

• Bring your transcripts

Tuesday class…

• First Project milestone: Empathy
  ✿ due next Tuesday

• No reading for next Tuesday