

First approaches to qualitative data analysis

I214 9 Oct 2008

Recap: Collecting (mostly qualitative) data

- ▶ **Observation**
 - ▶ Field notes: your own notes on what you see and think
 - ▶ Video, photography
- ▶ **Interviews**
- ▶ **Participant observation**
- ▶ **Self-reports**
 - ▶ Diary studies (people keep track of specific events x time)
 - ▶ Experience sampling (we periodically ask people what is currently happening)
 - ▶ Cultural probes
- ▶ **Document and artifact analysis**



What are we looking for?

- ▶ Answers to specific questions
- ▶ Problems, concerns that need further investigation
- ▶ Design requirements, ideas, problems that need to be solved
- ▶ Patterns, themes



Patterns

- ▶ Recurring issues, themes
- ▶ Typologies
 - ▶ Users
 - ▶ Uses/activities
- ▶ Correlations (vs causality)
 - ▶ Including group differences
- ▶ Temporal patterns
- ▶ Higher-order abstractions



Answers to specific questions

- ▶ Questions asked as part of the interview protocol
- ▶ Questions NOT asked
 - ▶ That arise later
 - ▶ That can't be *asked*



Correlations

- ▶ **May be either**
 - ▶ A priori questions
 - ▶ those that emerge from the data
- ▶ **Correlations vs causality**
- ▶ **Kinds of correlations of interest**
 - ▶ Group differences
 - ▶ Contextual factors



Possible correlation that arose in interviews

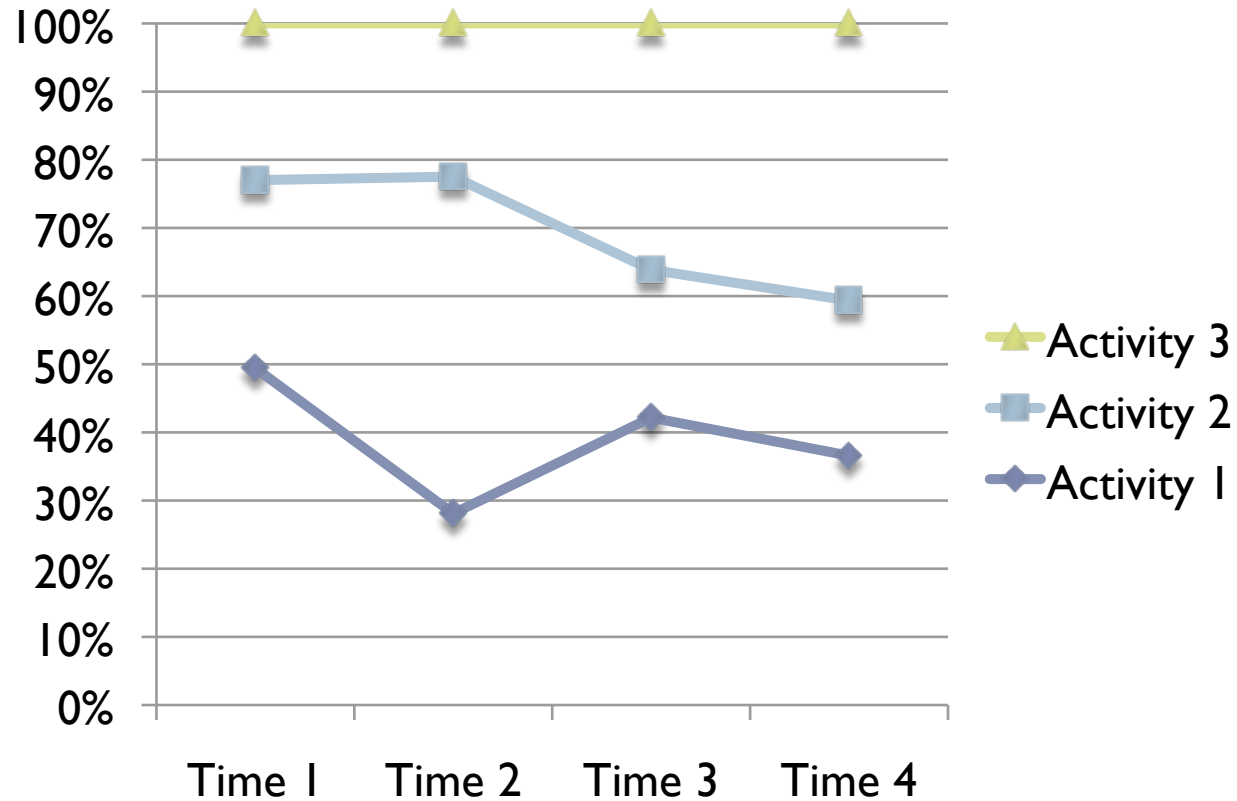
- ▶ Almost everyone who was serious about photography started young:
 - ▶ high school or (more often) before
 - ▶ Many took classes in high school
 - ▶ *Does doing photography, especially more seriously, correlate with early intro?*



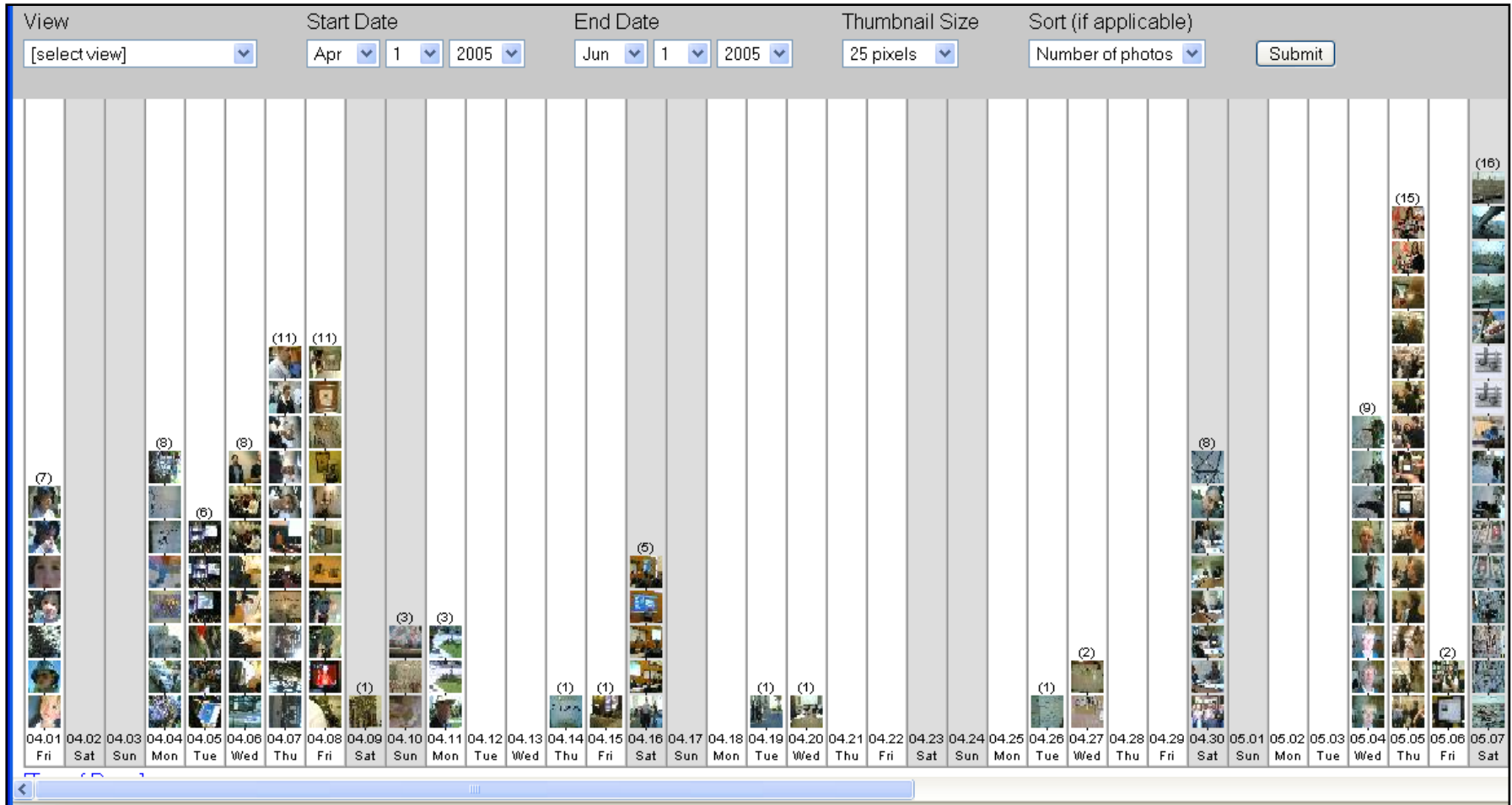
Group differences example: Parents and Flickr

- ▶ **Already knew parents differ from non-parents:**
 - ▶ Kinds of pics: Lots of pics of kids
 - ▶ Sharing demands: distant family, friends
 - ▶ Privacy: pics of kids online
- ▶ **New to us/we hadn't really thought about:**
 - ▶ Working vs non-working parents: time and access to computing
 - ▶ Location of computer at home: public vs private space
 - ▶ Time on computer is limited, with multiple demands
 - ▶ Working with computer while supervising kids
 - ▶ Computer as similar to TV as “babysitter”




Temporal Patterns



Photos over time

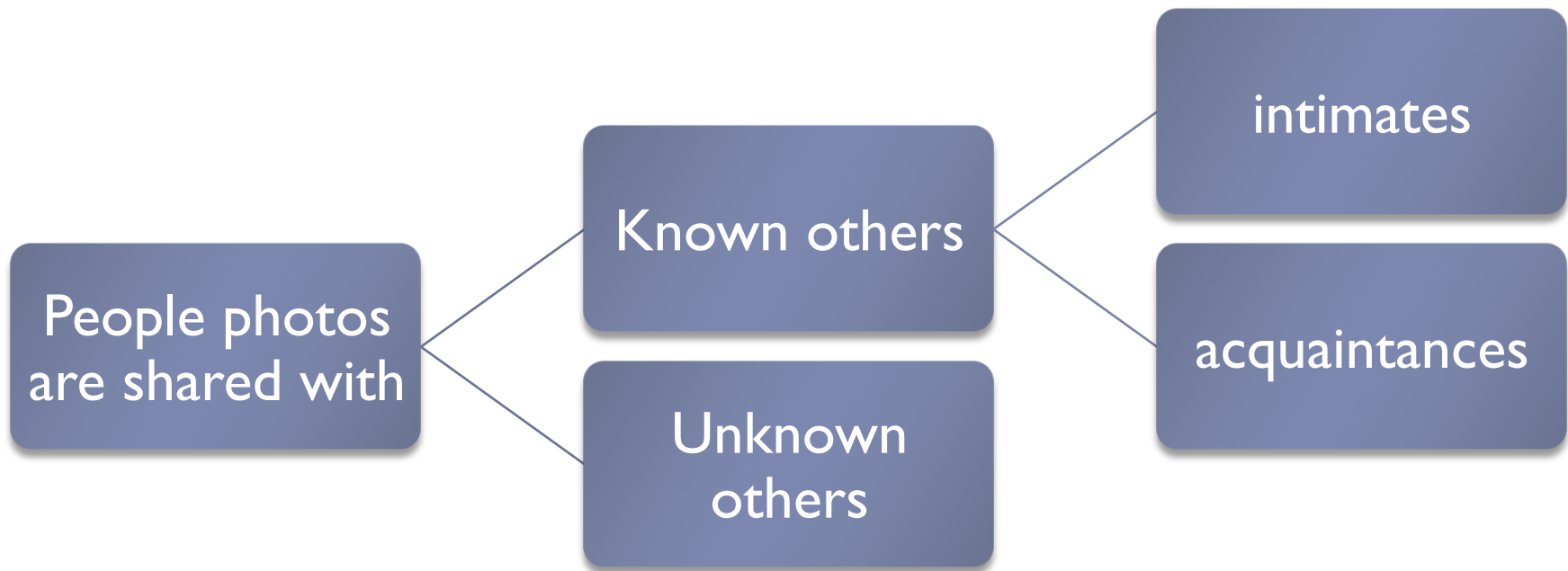


Typology: Users, Uses

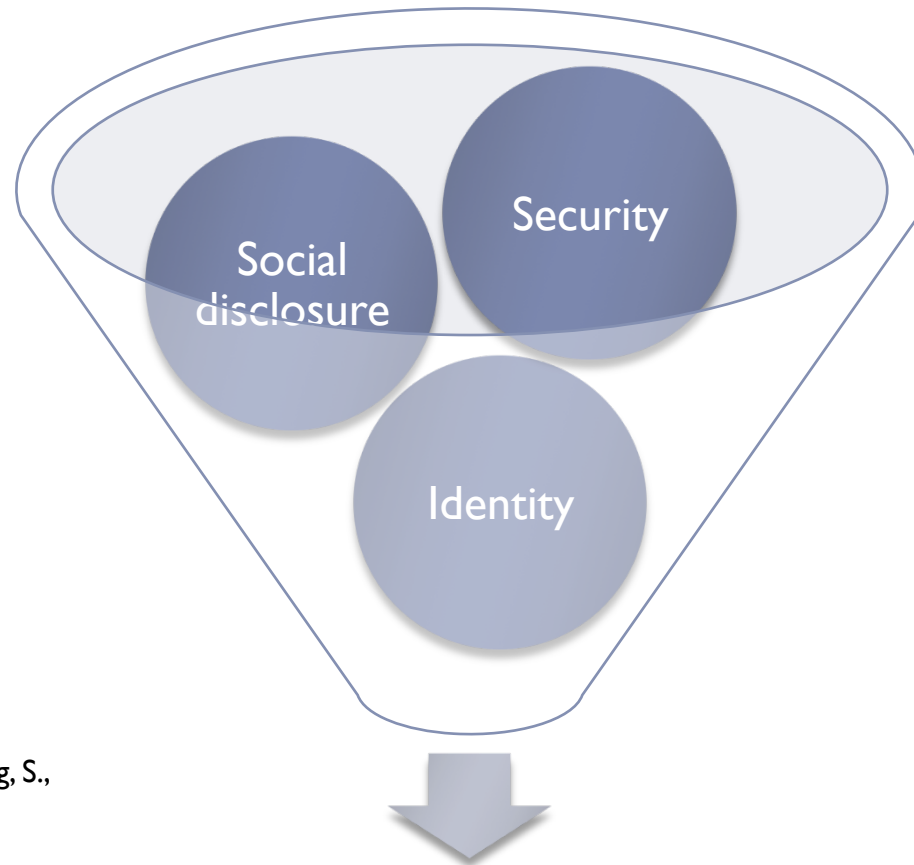
| Uses | Snapshot Photographers | Art Photographers |
|-----------|--|---|
| Snapshots |  |  |
| Art | |  |



Categories



Higher-order abstraction



Privacy on Flickr

Ahern, S., Eckles, D., Good, N. S., King, S., Naaman, M., and Nair, R. 2007. Over-exposed?: privacy patterns and considerations in online and mobile photo sharing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (San Jose, California, USA, April 28 - May 03, 2007). CHI '07. ACM, New York, NY, 357-366. DOI= <http://doi.acm.org/10.1145/1240624.1240683>

EX: Diary Study, Mobile Info Needs

- ▶ 20 people's mobile info needs x 2 weeks
- ▶ Questions:
 - ▶ Types of info needs (i.e., questions that users had)
 - ▶ Strategies and methods they used to meet them
 - ▶ Contextual factors that prompted each need & influenced how addressed
- ▶ Method: Diary study (ESM would need too high a sampling frequency); critical incidents
- ▶ Participants would send a “snippet” text message each time to special email address >> posted on web site
 - ▶ Reminded 5 times a day, every three hours
- ▶ At the end of the day, participants would log to describe event more fully

Sohn, T., Li, K. A., Griswold, W. G., and Hollan, J. D. 2008. A diary study of mobile information needs. In Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems (Florence, Italy, April 05 - 10, 2008). CHI '08. ACM, New York, NY, 433-442. DOI= <http://doi.acm.org/10.1145/1357054.1357125>

Questions answered for each incident

1. Where were you?
2. What were you doing?
3. What was your info need?
4. If you attempted to address it, how? If not, why not?
5. Could you have addressed your need by looking at your personal data (email, calendar, web browsing history, etc.)?



Quantitative results

- ▶ 421 diary entries
- ▶ Averaged 21.1 entries/person (min 7, max 45)
- ▶ No missing data (all snippets get full entry)



Analysis

- ▶ Taxonomy of information needs
- ▶ Factors involved in deciding when/whether to address need
- ▶ Needs address now, later, not at all
 - ▶ How addressed if now
 - ▶ If later, why
 - ▶ If never, why not
- ▶ How does mobile internet change behavior? (can't ask)
- ▶ What prompt information needs? (*Could context-aware computing help?*)
- ▶ Where does the information come from? (*What can we do to help provide it?*)

Taxonomy of information needs

| Need Category | Example | % of Total Diary Entries | # of Category Entries | # of Participants Reporting Category |
|------------------------|---|--------------------------|-----------------------|--------------------------------------|
| Trivia | "What did Bob Marley die of, and when?" | 18.5% | 78 | 17 |
| Directions | "Directions to Sammy's Pizza" | 13.3% | 56 | 17 |
| Point of Interest | "Where is the nearest library or bookstore?" | 12.4% | 52 | 17 |
| Friend Info | "Where are Sam and Trevor?" | 7.6% | 32 | 8 |
| Shopping | "How much does the Pantech phone cost on the AT&T website?" | 7.1% | 30 | 16 |
| Business Hours | "What time does the post office close?" | 6.9% | 29 | 15 |
| Personal Item | "What is my insurance coverage for cat scans?" | 6.4% | 27 | 12 |
| Schedule | "Is there an open date on my family calendar?" | 5.7% | 24 | 12 |
| Phone # | "What is the phone # for weight watchers?" | 5.7% | 24 | 13 |
| Traffic | "How far does the traffic extend?" | 4.5% | 19 | 8 |
| Sports/News/ Stocks | "Did the Miami Heat have any free agent acquisitions?" | 3.8% | 16 | 7 |
| Email | "Email update for work" | 2.6% | 11 | 3 |
| Movie Times | "Are Harry Potter tickets available tonight?" | 2.4% | 10 | 9 |
| Weather | "What will the weather be like this weekend?" | 1.4% | 6 | 4 |
| Travel | "Flight status of my Southwest flight" | 1.0% | 4 | 3 |
| Recipes | "Needed ingredients for hot and sour soup" | 0.7% | 3 | 3 |



Need categories (and exs)

Trivia

Directions

Point of interest

Friend info (where is Joe?)

Shopping (price of x)

Business hours

Personal item (my insurance)

Schedule (own, family)

Phone #

Traffic

Sports/news/stocks

Email (email update for work)(?)

Movie times

Weather

Travel (my flight)

Recipes



How can we group these?

Trivia

Directions

Point of interest

Friend info (where is Joe?)

Shopping (price of x)

Business hours

Personal item (my insurance)

Schedule (own, family)

Phone #

Traffic

Sports/news/stocks

Email (email update for work)(?)

Movie times

Weather

Travel (my flight)

Recipes

