

# **Research Methods In Human- Computer Interaction (2<sup>nd</sup> edition)**

Chapter 5 Surveys

2017/7/25

担当 森

# 1. Introduction to HCI Research

Topics of experimental design

## 2. Experimental Research

## 3. Experimental Design

## 4. Statistical Analysis

Useful for understanding  
“why” questions

## 5. Surveys

Two key approaches from sociology

## 6. Diaries

## 7. Case Studies

## 8. Interviews and Focus Groups

Research approaches popular  
in the social sciences

## 9. Ethnography

Methods of collecting research  
data

## 10. Usability Testing

## 11. Analyzing Qualitative Data

## 12. Automated Data Collection Methods

Issues that arise in working  
with human subjects

## 13. Measuring the Human

## 14. Online and Ubiquitous HCI Research

## 15. Working With Human Subjects

## 16. Working With Research Participants With Disabilities

# Chapter 5 Surveys

- 5.1 Introduction
- 5.2 Benefits and drawbacks of surveys
- 5.3 Goals and targeted users for survey research
- 5.4 Probabilistic sampling
- 5.5 Nonprobabilistic sampling
- 5.6 Developing survey questions
- 5.7 Overall survey structure
- 5.8 Existing surveys
- 5.9 Paper or online surveys? ← 今日はここまで m()m
- 5.10 Pilot testing the survey tool
- 5.11 Response rate
- 5.12 Data analysis
- 5.13 Summary

# 5.1 Introduction

- Survey とは
  - A well-defined and well-written set of questions to which an individual is asked to respond
  - HCIにかぎらず色々な分野で使われる手法
  - 被験者が個人で実施。民族誌学やフォーカスグループなどによる深い分析まではできない。
  - “big picture” を手軽に手に入れることができる。どのような問題があるのか、どのようなアクションをしているのか等々。
- Survey 手法が批判される理由:
  - 適切どころではなく、簡単だからという理由で使われることが多いため。
  - 例) 学生を対象にしたサーベイ結果を一般化してしたり、パイロットテストなしにサーベイを作ってしまうため。

# 5.1 Introduction

- HCIにはサーベイが最も適している研究プロジェクトがある
  - ユーザエクスペリエンスに対する attitude、awareness, intent, feedback やユーザの特性、時間が経過した前後の比較等の計測に適している
  - 精密な計測や単にあるインタフェースのユーザビリティの問題を特定する目的には向いていない
- サーベイは質問紙 (questionnaire) と同じ？
  - 質問紙は質問のリスト。サーベイはサンプリングやリマインダ、インセンティブなども含む方法論的なアプローチ全体を示す。
  - 質問紙よく練られたサーベイを構成する一要素 (Dillman)
  - 用語としてはあまり区別されずに使われることが多い(本書でも同様)

## 5.2 Benefits and drawbacks of surveys

- Benefits
  - 大勢の人から低コストでデータを収集することができる
  - ユーザー母集団の overview や snapshot を得ることが容易
  - 高度なツールを必要としない。E-mail やサーベイ用のweb サイト、紙の質問紙など。
  - IRBや human subject board からの了承を最も得やすい手法。(IRB については Chapter 15 参照→秋山さん担当)
- Drawbacks
  - 大規模な母集団から浅いデータを収集することに向いているが、深い分析には向いていない。
  - 参加者の自己管理(self-administrated)なので、面白い現象があっても follow-up question をしたり、元の質問を変更して詳細な質問をすることはできない。

# 5.2 Benefits and drawbacks of surveys

- Drawbacks つづき
  - 質問が明快な事実に基づく現象ではなく、利用パターンに関連している場合、バイアスを生じることがある。
  - 例)
    - 雰囲気 (mood) に関する質問は実際のイベントから時間が経過しているとバイアスを生じる可能性あり
    - このアプリを使った時にどう感じましたか？
    - 半年以内に e-commerce でいくら使いましたか？ あるソフトを使って特定のタスクを何回完了しましたか？
  - これらの質問に対する回答は overestimate か underestimate の可能性あり
  - このような場合、コンピュータで自動収集したデータが使えるならそのほうが望ましい。あるいはさらにサーベイを併用。
  - 調べたい対象が口頭でのコミュニケーションを好むカルチャーから来ている場合は民族誌学的アプローチが良い

# Researching photo tagging and sharing behaviors

- Nov et al. 2008
  - Flickr で写真にタグ付けした回数を調査
  - 回数はデータログから。他の RQ はサーベイを使った。
  - 少なくとも5このタグをつけたユーザをランダムサンプリングで抽出。
  - 該当ユーザにサーベイを実施した後、許可を得られたユーザに対してはデータ収集を実施。
- Kairam et al. 2016
  - Flickr でどのように共有する写真を選択するかを調査
  - Flickr ユーザに対して、調査の一環として写真にアクセスすることを許可することを条件に参加者を募集 (AMT)
  - 96の被験者に対して20の写真を選択。さらに10個をランダムにサンプリングして、個々の被験者毎に詳細に調査(サーベイ)。



# 5.3 Goals and targeted users for survey research

- ターゲット
  - 興味がある母集団を “target population” と呼ぶ。HCI の場合は targeted users。
  - Well-defined な母集団であれば母集団にふくまれる個人の数がわかる。
  - 誰がターゲットなのか、なぜそれらの人に関心があるのかを考えることが重要。
  - 誰でもサーベイに回答できるといえる場面は少いはず
  - 年齢、性別、ソフトウェアの経験、勤務先、障害の有無等々
  - サーベイデザインの世界では “targeted respondents” と “user population” や “inclusion criteria” は同様の意味で使われる。
  - Inclusion criteria: サーベイにおいてどのような人が含まれるべきかを詳細に記述する。

# 5.3 Goals and targeted users for survey research

- ターゲットを決めた後、どうやってそれらの人を集めるか？
- 一般的な社会学の研究では電話帳や電子メールアドレスリストを使う傾向にある。
  - 電話によるサーベイは最近はあまり使われない。電話勧誘の影響や固定電話の現象等。
- サーベイの回答者を決定するのにリストやリストの集合が使われることを “defining the population frame” と表現される
- HCI では一般的な集団よりは specific な母集団に関心があることが多い
- Inclusion criteria の例
  - ある特定の職業、ウェブサイト、リスト、特定の職業のソーシャルネットワーク、などに所属
  - 障害をもつ人であればそのような障害がある人の組織のディレクトリが適切
  - ソフトウェアのユーザであればソフトウェアのユーザとして登録した人のディレクトリが適切
  - SNS のグループを使うのも有用
- ただしこれらの方法単独ではうまくないことが多い
- サーベイの母集団を well-defined にするのが難しい場合、多様で複数のサブグループから回答を得ることを目指す。(5.5 節参照)

# 5.4 Probabilistic sampling

- 社会学における古典的なサーベイの利用は母集団に対して推定を行うこと
- すべてのメンバーに対してセンサス調査を行うのが最も精度が高いが、コストがかかるのでめったに行われない
  - センサスは大きな組織や政府関連組織がスポンサーとなっていることが多い
- 関心がある母集団が小さい場合（例えばたかだか数1000程度）、全数調査をしても良いかもしれない
- センサスの代わりにランダムサンプリングが使われることが多い
- 個人がどの程度の確率でサンプルに選ばれるかが決定している。
- サンプリングは個人ではなく、組織単位で行っても良い。

# 5.4.1 Stratification

- サンプルングは層化 (stratification) することができる場合がある。
- 母集団を部分母集団に分割し、それぞれの部分母集団から個別にサンプルを抽出する。
- 層化により、各々の部分母集団からサンプルが存在していることを明確にできる。
- 大学生をサンプルする際に、各学年にわけることができる。学年ごとの人数に差がある場合、層化でやれば各学年から同じ人数ずつサンプルを集めることができる。
- 例) 引っ越しをした後の友人関係に技術がどのような影響を与えるか
  - 引っ越しをした人を母集団とし、1/3 は近い引っ越し、2/3 は遠い引っ越しとして層化サンプルングした。これで距離に関する分析もできた。
- 森注: 層別サンプルングは部分母集団のサイズに大きな差がある場合に全体としてサンプルング誤差を小さくするために用いられる統計学の手法
  - サンプル数が小さいとマイノリティがサンプルされないことがあるので、比率を守ってサンプルングするように強制する。

## 5.4.2 Response size

- ランダムサンプリングが良い方法だとして、回答をいくつ集めればよいか？
- 統計学的に明確なしきい値というものはない。
- 確率的サンプリングの場合、信頼水準 (confidence level) をどうするか、誤差マージン (margin of errors) がどの程度かを定める必要がある。
- 単純なランダムサンプリング384個のサンプルサイズを得た場合  
⇒95%の信頼水準、±5%の誤差マージン。  
100回サーベイを実施した場合、95回は実際の割合から5%以内の割合を得る。
- 誤差マージンを±4%にする場合は600のサンプル、±3%にする場合は1067のサンプルが必要。
- サンプル数が母集団と比較して大きい場合(例えば5～10%)、誤差マージンは小さくなる。
- 有限母集団補正 (finite population correction) という方法もあるが、本書では取り扱わない。
- Muller et al. 2014 pp. 238-239 では HCI research でのサンプルサイズに関する説明がある

## 5.4.3 Errors

- ランダムサンプリングは理想的な手法であるが、誤差やバイアスもある。
- これらの要因に注意を払うことがサーベイで得た発見の精度や信頼性を高める。
- カバレッジ誤差: 関心のあるメンバーが等しくサンプルされない場合 (リストに乗っていない人など)
- 計測誤差: 質問がちゃんと書かれていなかったりバイアスを持っている場合
- 回答なし誤差: demography で大きな差異がある場合

# 5.5 Nonprobabilistic sampling

- 5.4 での probabilistic sampling の目的は母集団推定であったが、HCIでのゴールは必ずしも母集団推定ではない  
⇒nonprobabilistic sampling が使われることが多い。  
決まった手順やお作法があるわけではない。
- HIV感染患者を対象にした場合、患者情報の秘匿性から任意のサンプリングフレームワークを当てはめたり厳密なランダムサンプリングを実施することは困難。  
⇒しかし妥当な方法で nonprobabilistic sampling を実施することはできる。
- Nonprobabilistic sampling の例:
  - Opt-in panels, self-selected surveys (website や SNS のリンクをクリック), snowball recruiting (回答者が他の回答者候補を選ぶ)

# 5.5 Nonprobabilistic sampling

- サンプルング技術の適用の仕方は分野毎にスタンダードが違う
  - 社会学や統計学の人々は厳密なランダムサンプリングでないとサーベイは妥当ではない考える
  - HCI では確率的サンプリングではない方法を長らく使ってきており、それが妥当で許容可能だと考えられている。
  - 収集されるデータの違い: センサス vs. HCI 実験
- HCI 研究では他の研究手法とともにサーベイがきわめてよく使われる
  - かならずしもサンプルが代表的だとは主張しない
  - CHI 2016の論文では1000以上の回答があるサーベイを使った研究もあれば、diary study や interview、usability testing, long analysis など研究手法と小規模なサーベイを組み合わせた研究もある



# 5.5.1 Demographic data

- Demographic を収集・分析することで回答者が多様で複数のセクションにまたがり、もしあればベースラインとなるデータに対して代表的であることを示す。
- 例) 年齢、性別、学歴、職種、コンピュータの利用経験等
- ランダムサンプリングによる母集団推定とはまったく異なるが、何も示さないよりは良い
- Harbach et al (2016): 8カ国にまたがる8000の回答者によるサーベイ。(smartphone locking)  
→ demographic と国ごとの文化に関する情報

## 5.5.2 Oversampling

- 良く定義されたユーザリストがなく、ランダムサンプリングもできない場合、回答者数が重要となる
- Nonprobabilistic sampling では20のサーベイ回答では不十分であろう(demographic data があっても偏りがある)
- 推定される母集団に対して十分に大きな回答数があれば妥当だとみなされる (oversampling)
- 例えば関心がある母集団のサイズが5000であれば500の回答数は大きいと考えられるが、母集団のサイズが国レベルだとすると必ずしもそうではない
- Sue and Ritter 2007 ではどのようなサーベイリサーチであっても最低30回答がベースラインと述べている

## 5.5.3 Random sampling of usage, not users

- ユーザではなく、利用 (usage) をランダムランブリングする
- ユーザはあるウェブページに10回アクセスするたびにサーベイに回答することを求められる (pop-up window などで)
- Intercept sampling としても知られる (Sue and Ritter, 2007)
- Usage については正しくとれるが、そのウェブページにアクセスする人が over-represented となり、アクセスしない人は under-represented となる。

# 5.5.4 Self-selected surveys

- Self-selected survey: ある website にアクセスした人は全員サーベイに回答することを求められる
- ある集団の人々が選ばれたというよりはすべての人に参加を求めているのに近い
- 回答数とデモグラフィックデータが重要
- ジョージア工科大学のバナー広告を使った例が有名
  - [http://www.cc.gatech.edu/gvu/user\\_surveys/](http://www.cc.gatech.edu/gvu/user_surveys/)
  - <https://scholar.google.com/citations?user=jYHJAagAAAAJ&hl=en>
- 確率的サンプリングが現実的ではないことが多いので、そういう時に使う
  - <http://webaim.org/projects/screenreadersurvey6/>
  - <http://chrishofstader.com/screen-reader-market-figures-my-analysis-of-webaim-survey-6/>
- 新しい母集団や新しい利用の現象などを調べるのには self-selected survey が適している
  - サーベイは social networking group, メーリングリスト、あるいはそのような母集団が集まるチャットルームなどで実施。

# 5.5.5 Uninvestigated populations

- 詳細に調査されていないユーザ集団がある
  - 例) 認知症患者
- 仮説や実験デザイン、良く構成された日記などを作るためのベースラインとなるデータがない。
- ある国における特定の疾患を持つ人の母集団推定はあったとしても、それらの疾患をもつ人の内、何人がコンピュータ技術を使っているかというデータはない
  - ⇒ランダムサンプリングは適切ではない
- そのようなサーベイのゴールはベースラインデータを確立することにある
  - ⇒できるだけ多くの回答者を収集(ダウン症患者のコンピューター利用の例)
- Snowball sampling が有益なケースがある。

# 5.6 Developing survey questions

- サーベイを使うゴールと戦略が決まったら次はサーベイツールを開発するステップ
- サーベイツールは質問紙とも呼ばれる。
- よく書かれたバイアスがない質問を書くことが重要
- Self-administrated なことが多いので、意図しない回答がないようにする。
- サーベイには2つの構造がある
  - 個々の質問の構造
  - サーベイ全体の構造
- 個々の質問は3つ方法で構成できる
  - Open-ended
  - Close-ended with ordered response categories
  - Close-ended with unordered response categories

# 5.6.1 Open-ended questions

- 回答者が柔軟に回答できるので、現状を理解するのに有用
- 質問は慎重に記述する必要あり
- 例)
  - *Why did you stop using the Banjee Software product?*  
⇒原因についての情報がない⇒回答者は深く考えないといけない  
⇒回答しないか、“I didn’t like the software” のような回答をする
  - *How did you feel about the usability (ease of use) of the Banjee software?  
Did the Banjee software allow you to complete the tasks you wanted to complete?*  
⇒より具体的なので“I didn’t like..”とは答えないが、単純に Yes/No で答えてしまうかもしれない
  - *What barriers did you face, in attempting to use the Banjee software to complete your tasks?*  
⇒“none” とこたえるかもしれないが、これまでに体験した問題について慎重に考えることをうながす

## 5.6.2 Closed-ended questions

- Ordered response: 選択肢の数が与えられており、その順番が論理的。
  - 例) excellent to poor, strongly agree to strongly disagree
  - Fig 5.1
- Unordered response: 選択肢に論理的な順番がないケース。複数の選択肢を選択させることが多い
  - Web インタフェースのときは要注意 (option button ではなく checkbox)
  - Fig 5.2, 5.3



# 5.6.3 Common problems with survey questions

- サーベイの質問に下記のような問題がないかをチェックすべき (Babbie, 1990)
- 二連(double-barreled)の質問がないか。
  - 例) how long have you used the word processing software and which advanced features have you used?
  - これらは分けるべき
- 否定語が入っていないか。
  - 例) Do you agree that the e-mail software is not easy to use?
  - 混乱を招くので避ける
- バイアスがある言葉の利用はないか。
  - 例) Don't you agree that .. や、著名な人や組織の意見を先に提示してしまう
- “Hot-button” words を使っていないか。
  - Liberal, conservative, abortion, terrorism など

# 5.7 Overall survey structure

- サーベイはどのようなフォーマットであっても instructions から始まらなければならない
  - すべての質問に回答する必要があるか、名前やコンタクト先を書く必要はあるのか
  - 誰が回答者なのかを明示的に書いておくと良い (you must be aged 65 years or older)
  - いくつかのセクションからなる場合、区切りを明確にする
  - 回答者はサーベイをどの順番でどこをみていけば良いのかを理解する必要がある。
  - フォーマット (paper, e-mail, web) が違えば異なる情報やインストラクションを提供する必要がある
- サーベイのレイアウトは重要
  - 紙の場合は余白をとることが重要 (情報量に圧倒されないように)
  - 30ページでも、同じ情報を2ページに詰め込んでも回答者は不安に感じる
  - 白い紙、十分に大きなフォント、標準的なテキストを使うべき
  - 片側ホチキスよりはブックレット印刷が好ましい
  - 略語は混乱を避けるために使わない
  - Webの場合、さらなる情報を見るのにポップアップウィンドウを使うこともある

# 5.7 Overall survey structure

- サーベイの質問は研究の文脈上意味をなすのであればどのような順番でも良い
- しかし、似たトピックやアイデアに関する質問はまとめた方が良い (Dillman, 2000)  
⇒ 回答者の認知的負荷を下げることでトピックに深く思考することができる
- 面白い質問を先に出すことでモチベーションを高める
- 質問数が多すぎると回答率は下がる。必要な質問はすべて採用すべきであるが、個人がサーベイに回答する時間に関しては別途検討する必要がある。
- 最も簡単なサーベイはすべての回答者がすべての質問に回答する方式であるが、すべての回答者にあてはまらない質問もある  
⇒ Fig 5.4 のような例 (indented box) でやると良い  
web: If you answered no, please click here to move on to the next section

## 5.8 Existing surveys

- HCIコミュニティで検証されてきた既存のサーベイがあることを知っておくと良い
- Table 5.1 Survey Tools in HCI
  - QUIS: <http://www.lap.umd.edu/quis/>
  - SUMI: <http://sumi.uxp.ie/>
  - WAMMI: <http://wammi.uxp.ie/>
- 同じサーベイならそれらを使うことも検討
- 最小限の修正を加えて使う

# 5.9 Paper of online surveys?

- Paper, e-mail, web のそれぞれで pros/cons あり
  - 視覚障害者は paper ができない、ネットワーク使えない人は e-mail/web ができない等々
- 研究者が関心のある母集団にアクセスするためのベストな方法に応じて選択が決まることが多い
  - 個別に訪問する場合⇒紙が良い
  - E-mail list がある場合 ⇒ e-mail が良い
- Web-based サーベイ
  - SurveyMonkey, SurveyPlanet, FreeOnlineForms.com, Google Forms
  - 最近是最もよく使われる手法
- 紙と web のコンビネーションが使われることもある。回答者に選ばせる⇒回答者の好みにあわせることで回答率向上につながる(回答方法によるバイアスが生じないように注意)
- 複数言語で質問を作る⇒自動翻訳を使っては駄目

# 5.9 Paper of online surveys?

- 電子的なサーベイはコストが低く、早く完了できるメリットがある
- 電子的なサーベイの回答は信頼できるか？
- 紙と電子でどちらが正直になるか(あるいは嘘をつくか)という実証実験結果はないが、bad news に関するトピックでは電子的な方法の方が正直になる
- 匿名回答であれば self-disclosure のレベルがあがる  
⇒ web の方が e-mail よりも良いケース
- Interview よりも self-administrated survey の方が正直に回答する
- 伝統的な紙のサーベイでは回答者は informed consent form に署名する  
詳細は Chapter 15

## 5.10 Pilot testing the survey tool

- 次回にて