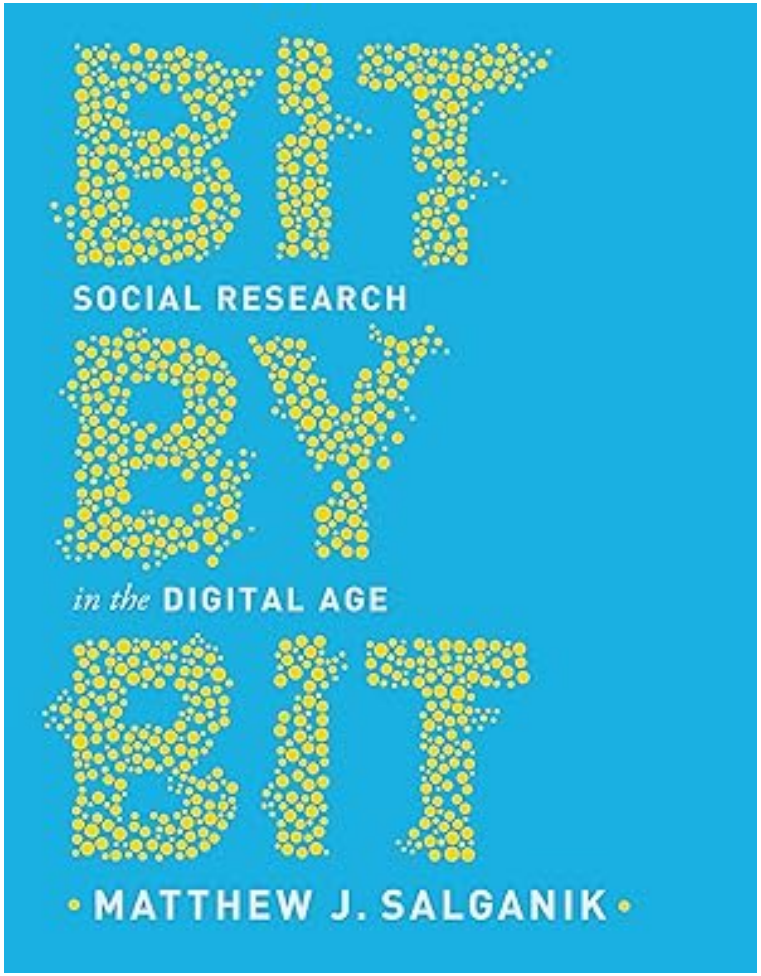
Socio-Informatics 348

Intro to Computational Social Science

Dr Lisa Martin

Department of Information Science
Stellenbosch University

Today's Reading



Bit by Bit, Chapter 6: Ethics

The Dual Problem of Ethics

1. **Visible Problem:** Some researchers have violated privacy or conducted unethical experiments
2. **Hidden Problem:** Ethical uncertainty has a *chilling effect*, preventing important research

Three Notable Ethical Controversies

Example 1: Emotional Contagion

What Happened

- 700,000 Facebook users in an experiment (January 2012)
- Manipulated News Feed to test emotional contagion
- Groups: negativity-reduced, positivity-reduced, two controls
- No consent beyond standard terms of service
- No meaningful third-party ethical review

Outcome

- Enormous public outcry
- Journal issued "editorial expression of concern"
- Led to creation of ethical review process at Facebook
- May have driven research "into the shadows"

Example 2: Tastes, Ties, and Time

What Happened

- Scraped Facebook profiles of Harvard Class of 2009
- Merged with university records (majors, housing)
- Used for research on social network formation
- Data shared with other researchers
- Students did not provide informed consent

Outcome

- School identity deduced within days
- Accused of "failure to adhere to ethical research standards"
- Dataset removed from Internet
- No longer available to researchers

Example 3: Encore

What Happened

- System to measure Internet censorship (March 2014)
- Website owners installed code snippet
- Visitors' browsers secretly tested access to potentially blocked sites
- IRB declined review (not "human subjects research")

Ethical Concerns

- People could be at risk if accessing sensitive websites
- No consent from participants
- Modified to only test Facebook, Twitter, YouTube
- Paper published with unprecedented "signing statement"

Why Digital Age Research Raises New Issues

Increased Power

Researchers can now:

- Observe behavior without consent or awareness
- Enroll people in experiments without their knowledge
- Do this at massive scale (millions of people)

Unclear Guidelines

- Rules changing slower than capabilities
- Norms around privacy still being debated
- Overlapping contexts (university vs. company)

Two Perspectives on Data

Master Database

- Exciting research possibilities
- Better understanding of society
- Improved services and policies

Database of Ruin

- Potential for unethical use
- Unanticipated secondary use
- Historical examples (WWII)

Key Insight

The same database can be *both* a powerful research tool and a potential source of harm

Foundation: Two Key Reports

Belmont Report (1979)

- Created in response to ethical failures (Tuskegee Syphilis Study)
- Four years of deliberation by national commission
- Intellectual basis for Common Rule
- Three principles: Respect for Persons, Beneficence, Justice

Menlo Report (2011)

- Response to ethical issues in computer security research
- Applies ethics to information and communication technologies (ICT)
- Reaffirms Belmont principles
- Adds fourth principle: Respect for Law and Public Interest

The Four Principles

1. **Respect for Persons**
   Treating people as autonomous and honouring their wishes
2. **Beneficence**
   Understanding and improving the risk/benefit profile
3. **Justice**
   Ensuring fair distribution of risks and benefits
4. **Respect for Law and Public Interest**
   Beyond participants to all relevant stakeholders

Principle 1: Respect for Persons

Two Components

1. Individuals should be treated as **autonomous**
2. Those with diminished autonomy entitled to **additional protections**

In Practice

- Informed consent (when possible)
- Participants control their participation
- Not what researcher thinks is best—what participant wants

Application to Examples

All three studies did things to participants without consent:

- Emotional Contagion: enrolled in experiment
- Tastes, Ties, Time: used their data
- Encore: used their computers for measurements

Principle 2: Beneficence

Two Components

1. **Do not harm**
2. **Maximize benefits, minimize harms**

In Practice: Two-Step Process

1. **Risk/benefit analysis** (technical)
   - Understand probability and severity of adverse events
   - Improve study design to reduce risks
2. **Ethical assessment** (values-based)
   - Does the study strike appropriate balance?
   - Some risks render research impermissible regardless of benefits

Beneficence: Suggested Improvements

Emotional Contagion

- Screen out people under 18
- Screen out vulnerable individuals
- Use efficient statistical methods (smaller sample)
- Monitor participants for harm

Tastes, Ties, and Time

- Additional safeguards when releasing data
- Better protection against re-identification

Encore

- Minimize number of risky requests
- Exclude participants in danger from repressive governments

Principle 3: Justice

Core Question

Are risks and benefits distributed fairly?

Historical Concern: Protection

- Vulnerable people exploited in research
- Poor bore burdens, rich reaped benefits
- Need to protect vulnerable populations

Modern Concern: Access

- Around 1990, focus shifted to inclusion
- Children, women, minorities should benefit from research
- Being excluded from research is also unjust

Justice: Application to Examples

Encore

Tension:

- Beneficence suggests excluding people in repressive countries
- Justice suggests they should participate and benefit

Emotional Contagion

Well-aligned with Justice:

- Participants = random sample of Facebook users
- Those bearing burden are those who will benefit

Tastes, Ties, and Time

- Harvard students bore burden
- Society as whole benefited
- No compensation to participants

Principle 4: Respect for Law and Public Interest

Extends Beneficence

- Beyond specific participants
- Include all relevant stakeholders
- Explicitly incorporate law

Two Components

1. **Compliance**
   - Identify and obey relevant laws, contracts, terms of service
   - May be situations where violation is permissible
2. **Transparency-based accountability**
   - Clear about goals, methods, results
   - Take responsibility for actions
   - Prevent secret research
   - Enable public input

Four Challenging Areas

1. **Informed Consent**
   Some form of consent for most research
2. **Informational Risk**
   Understanding and managing data risks
3. **Privacy**
   Appropriate flow of information
4. **Decision Making Under Uncertainty**
   Proceeding when risks are unknown

Area 1: Informed Consent

Simple Rule (Wrong)

"Informed consent for everything"

Better Rule

"Some form of consent for most research"

- Informed consent is neither necessary nor sufficient for ethics
- Respect for Persons is just one principle
- Must balance with Beneficence, Justice, and Respect for Law

Example

Field experiments on discrimination: no consent, but ethically permissible under specific conditions (Eg. job applications)

When Informed Consent is Difficult

Three Reasons

1. **Increases risk to participants**
   Solution: Public information, opt-out, group consent
2. **Prior consent compromises scientific value**
   Solution: Debriefing after study
3. **Logistically impractical**
   Solution: Consent from sample of participants

Warning

Research without any consent is in a gray area—be very careful and prepared to defend your decisions

Area 2: Informational Risk

Definition

Potential for harm from disclosure of information
Economic (losing job), Social (embarrassment), Psychological (depression), Criminal (arrest)

Key Challenge

- Most common risk in social research
- Has increased dramatically in digital age
- Hardest risk to understand and manage
- "Anonymization" is deeply flawed (see examples)

Key Lessons

1. All data are potentially identifiable
2. All data are potentially sensitive

Area 3: Privacy

Public/Private Dichotomy (Too Simple)

If information is publicly accessible $\rightarrow$ OK to use?

Problem: This approach is too blunt

Better Approach: Contextual Integrity

Privacy = Right to appropriate flow of information

Context-Relative Informational Norms

Three parameters:

1. **Actors** (subject, sender, recipient)
2. **Attributes** (types of information)
3. **Transmission principles** (constraints on flow)

Key Question

Does this use violate context-relative informational norms?

Privacy: Two Perspectives

Harm-Based View

- Privacy violations matter if they cause harm
- Focus on consequences
- More consequentialist

Rights-Based View

- Privacy violations are harms in themselves
- Focus on dignity and respect
- More deontological

Example: Secret Surveillance

- Harm-based: Only problem if detected or used harmfully
- Rights-based: Violation occurs regardless of detection

Area 4: Making Decisions Under Uncertainty

The Problem

- Often must decide without complete information
- What's the probability of harm?
- How severe could harm be?
- Digital age = less experience, more complexity

Precautionary Principle (Problematic)

"Better safe than sorry"

- Can cause harm by preventing research
- Chilling effect on important studies
- Focuses only on risks of action, not inaction
- **There is no risk-free approach**

Tools for Decision Making Under Uncertainty

1. Minimal Risk Standard

- Compare study risk to everyday risks (sports, driving)
- Easier to assess relative risk than absolute risk
- Example: Is experimental News Feed like normal News Feed?

2. Power Analysis

- Calculate minimum sample size needed
- **New emphasis:** Make sure study isn't too big
- Minimize risk by minimizing participants
- Digital age: Avoid over-powered studies

More Tools for Uncertainty

3. Ethical-Response Surveys

Ask potential participants two questions:

- Would you want someone you care about to participate?
- Should researchers be allowed to proceed?

Benefits: Get perspectives before study, see public reaction

4. Staged Trials

- Like drug development: Phase I (safety) $\rightarrow$ Phase II (efficacy) $\rightarrow$ RCT
- Start with small, safety-focused studies
- Example: Encore could start in countries with rule of law
- Gradually expand after assessing safety

Key message: Uncertainty need not lead to inaction