

Introduction to Digital and online Research and the ethical issues

EUDACT 2015

Outline

- Development of digital research (lecture)
 - Ethical issues – discussion
-
- How has research into the digital world deployed different methods to collect and analyse data?
 - How have ‘digital research’ methods developed?

Type of approaches

- Research on online activities, interactions, communities
 - Research on impact of the online on the 'real'
- Merger of 'virtual' and 'real': the 'digital world'
1. Research the 'real' world through the digital
 - 'Digital social research'
 2. Research on the 'digital world' – a place where the digital is largely domesticated by the population.
 - Reflexive approach to the digital in the world.

Risks

A taken for grantedness of the 'digital' – a blindness to the changing technological, cultural, governance forms and practice

- E.g. studies of Facebook – although Facebook as a thing has changed significantly over the years; same also for computers, internet, mobile, ecommerce etc
- Concepts such as the 'Born digital'

When is it appropriate to step 'outside', when is it appropriate to take the digital for granted?

Questions...

Is the online world different, and needs
new techniques and approaches?

No, but...

Does the online world offer new
techniques of research?

Yes, but...

Main challenges to research ethics

- Blurring of Public and Private (privacy)
- 'Data' v. 'people' (e.g. copyright, 'human subjects')
- Mass observation (Surveillance)
- Mass experimentation (Manipulation)

Disciplinary and organisation shaping of digital research

- Disciplinary traditions and ethical codes shape the adoption or transformation of research approaches
 - Epistemological, ontological differences
 - E.g. micro v. macro, experimental, descriptive, role of theory etc
- Organisational context and goals e.g.
 - Knowledge creation (Science)
 - Commerce
 - Management
 - Policy
- Confidence with computers and the Internet of individuals and institutions

Research themes 1: about the digital world

Communications: How *people interact in text environments*, following phone and face to face studies.

- Shift from deficiency model to (i.e. face to face is best) to neutral model
- Socio-psychology: change *in group dynamics* (consensus, polarisation), the role of offline status etc (Lea et al, Sproll, Kielser et al etc) in existing organisations:
- Method: observation of email exchanges, bulletin boards; Experiments in group dynamics

Internet and Digital Research

Digital social research has been conducted since the earliest days of online systems –

- Bulletin Boards,
- ‘The Well’ ,
- Corporate email systems
- Multi-user online games (MUDS)

About the digital world : the age of virtuality

- The development of ‘online community’ (e.g. Rhiengold),
 - how they work, and the value people gain from participation, the dynamics of governance: used *participant observation*
- *Why* and *How* individuals use online environments
 - E.g. Turkle: Participation, and interviews
 - Tradition of online ethnography and participant observation in studying online game environments; chatrooms etc: *How they work*.
 - Exploration of on impact on ‘real’ life of individuals

Themes 2: Impact Studies

- What is the Impact of people interacting online on the 'real world'?
 - E.g. Putnam's challenge that electronic media is undermining community
- Wellman et al (Toronto): studies of communities using online services *in situ*
 - Mixed methods: Interviews; surveys of sociality and interactions; reading bulletin boards, longitudinal etc
- The rise of social network analysis: how does 'internet in community' change the community
 - Indicators of community participation, social capital etc

Today

- Online is as much real as 'offline'
- Ethnographies and many other methods are increasingly blind to the boundaries
- E.g. Garcia et al 2009; Miller and Horst (eds) 2012; Miller and Slater 2000

Impact 2

How does the internet impact on:

- Business (markets, opportunities)
- Politics (power)
- Policy (choices, concerns)
- Society (gender, inclusion, morals, family etc)
- Science, medicine, war, etc

Study emerging Practices using digital technologies and services –

- ‘inside’ – emerging practices – emergent or/and directed.
- ‘outside’ – incorporation or domestication of digital practices into wider society, and shaping of society

E-commerce Research

- Consumer behaviour in the digital world
 - How to get people to buy online?
 - Optimisation of websites (UX)
 - Understanding of consumer
 - *Personalisation of services (ads, marketing, etc)*
- Experimental methods on user behaviour e.g. Amazon , Facebook (Kramer et al)
- Dataveillance (Facebook, Google, NSA etc)
- Google adwords – opens up practical experimental research to advertisers

Slowly, administrative data is becoming to be seen as valuable for research and development

Surveillance as Research

- Observation
 - In Public Places
 - In Private Places
- Observation
 - of activities
 - of transactions
 - of interactions
 - of relationships
 - of communication

Observation in the Digital World

- Video surveillance
 - Public
 - Private – in homes, work places
- Data observations
 - Financial: Bank records, Sales, Tax
 - Communication: Emails, tweets, bulletin board posts etc
 - Geographical: Transport, phone tracking
 - Search use
- Population v. Individual; mass v. targeted.

Surveys in the digital world

- Quantitative surveys are difficult and expensive (Eurobarometer questions cost 1m euros each approx).
- Ad hoc Longitudinal studies very difficult
 - Except most expensive Panel studies
- Online reduces cost, and can now reach much of the population
 - Lots of work on how people answer online surveys

But:

- Engagement is very difficult
- Many internet panels are made up of 'professional' respondents (Week 5)

Clickstream Data

- Panels of individuals who volunteer to have their internet use monitored on their computers
 - Commercial services can provide this data on payment
 - The individuals are profiled
- What do people look at, when, how etc
 - Example of trying to track music piracy
- Compare with server side logging – single service data, P2P monitoring, or honey pots to attract users etc

Big Data

(read Rob Kitchin, Thomas Davenport for example)

- Messy real world data – but there must be some value in it!
 - Phone companies mining data to eliminate or keep customers
 - Banks, ecommerce, tax – detecting fraud
 - Video analysis - security
 - Sensors – transport tracking by ticketing

Very hard and expensive to do on full data, in real time. Google works only on small samples.

- Potential for natural experiments

Big Data challenges

1. Beyond conventional analysis – new techniques and skills:

e.g. qualitative research using vast data sets

- Analysing natural language by machine
 - Compare with machine translation
- If it is not collected it is invisible

2. Ethics - Massive and intrusive reuse of data

Crowdsourcing and Citizen Science (CSCS)

- In Surveys, social monitoring etc – the subjects are the participants
- IN CSCS Participants are part of the research process :
 - Collecting Data
 - Analysing data/producing evidence/visualising
 - Using evidence
- Paid and voluntary
- Organisational tools
- Curation of the crowd
- Engagement
- Quality Assurance Tools
- Legitimacy

Mobile Research: the new frontier

- Direct to individuals
- Surveys (limited by format and context)
- Ethnographies
- Qualitative
- Event triggered
- Location tracking, app use and data logging
- Mobile a data collection tool (crowdsourcing)

Week 5

Big Data?

- Science 'Big data' – e.g. from astronomical observations, genomics
- Volume; lack of structure; Realtime high flow; multiple types
- Data sets not collected for research
 - the outputs of administrative (bureaucratic or technological) systems. Tax authorities, Facebook logs, Phone logs
- Not Data collected with a methodology design to answer a research question,

Serious Games and Models

- Much research is related to informing decision making
 - Complex issues are hard to understand e.g. tradeoffs.
 - *How does one find out what people who do, what priorities they would give, if they don't understand what they are being asked about?*
 - Research is used to build a model of the world – a theory
 - A simplified model can be used to communicate the main trade-offs and issues
- = A GAME
- e.g. agriculture, energy

Serious games and models

- By playing, people learn about the issues, and then can be surveyed or interviewed normally
- The use of the game – the ‘analytics’ can also be used to understand about understanding, and observe choices.
- Can be used with Policy makers and with general or targeted publics
- Can introduce the option to *change the model*

Summary: Challenges for Practice

Techniques
Skills = People
Organisation
Standards
Legitimacy
Consent and Privacy
Access
Law
Reuse (open data)

Ethics Reading

- Facebook example (Zimmer 2008)
- SAGE handbook paper (2008)
- Professional guidelines
- Other reading

Observations and Comments Please

Main challenges to research ethics:

Discuss

- Blurring of Public and Private (privacy)
- 'Data' v. 'people'
- Mass observation (Surveillance)
- Mass experimentation (Manipulation)

Ethics issues

- Harm v. Social value
- Dignity
- Confidentiality
- Informed Consent
- Observer role
- Public or private? (Tweets, wikileaks etc)
- Deception
- Reuse of data/ open Data

Homework and Assessment

- Aim: help you understand ethical issues associated with digital research.
- Write a short summary of a particular ethical issue
- Write a summary of the ethical issues associated with a particular research approach
 - 20%

Next Week

Open Data

Look for examples in your work or
field