The Ethics of Internet-based and Social Media Research

Introduction and aims:
The University Research Ethics Committee (UREC), in collaboration with the Digital Society Network and the Information School’s Digital Societies Research Group, held a workshop on 14th July 2016 focusing on the ethical challenges of internet-based research, with a particular focus on research using social media.

Research in this area has been developing rapidly in recent years, and brings with it a range of ethical issues, including challenges in relation to what constitutes personal and public data, when and how consent can/should be obtained, and how collected data can and should be used. The UREC’s review of the University’s Ethics Policy during 2015/16 brought these issues to the fore, with numerous researchers reporting that they would value guidance in this area.

This event aimed to bring together researchers with a vested interest from across the University, to discuss and debate the ethical challenges of internet and social media research, and to contribute to the development of University policy and guidance on this area.

The workshop’s objectives were to:

1. Explore the ethical challenges of undertaking research using the internet, and in particular social media
2. Provide a forum in which to discuss and debate how to tackle these issues
3. Contribute to the development of University policy and guidance that will support researchers in undertaking research of this type.

The programme:
The workshop was opened by Professor Peter Bath, Chair of the University Research Ethics Committee (UREC), who gave a short background and introduction to the event. Presentations were then given by Professor Helen Kennedy, Professor Digital Society in the Department of Sociological Studies and Director of the Digital Society Network, and Dr Natasha Whiteman, Lecturer in the Department of Media and Communications at the University of Leicester, and author of ‘Undoing Ethics – Rethinking Practice in Online Research’ (2012).
Presentation slides are available at http://www.sheffield.ac.uk/ris/other/gov-ethics/ethicspolicy/educationresources/social_media_workshop_july_16 and some additional points (i.e. not duplicating information given in the slides) from Professor Kennedy and Dr Whiteman's talks are provided later in this report.

The presentations were followed by discussion sessions, in which attendees were split into 3 large facilitated groups, and then split again into smaller groups to discuss one of four scenarios presenting ethical challenges in social media research. Each small group then fed the key points of their discussions back to the large groups, and further discussion took place. Key points from all of the discussions were recorded and are presented in the next few pages of this report.

The event was rounded off with a panel discussion and question & answer session, chaired by Professor Bath and involving Professor Kennedy, Dr Whiteman, and Dr Jo Bates of the Information School’s Digital Societies Research Group, who was one of the developers of the workshop and who provided the discussion scenarios. Key points from this session are presented at the end of this report.

Additional points from Professor Kennedy's presentation:

- How do we define research participants? The Association of Internet Researchers’ guidelines state that it may not always be clear if people are involved.

- Are people OK with their data being researched? Public opinion is that this is not OK, according to various studies.

- What do people think is a private space? This is key, not a technical definition of public versus private.

- Do researchers have a responsibility to do research on issues of key public interest?

- ‘Masking’ is a possible way of dealing with confidentiality issues (i.e. changing the wording of messages) BUT Twitter's Terms and Conditions state that Tweets must be given in their original form and attributed to the individual who posted the Tweet.

- Anonymity is particularly difficult to deal with when referring to images.

- Proper informed consent becomes very difficult or impossible when dealing with social media data due to the large scale – people are rarely aware of their participation.

- Do the social media platform’s Terms and Conditions constitute informed consent? E.g. Twitter does explicitly state data will be used for research, but how many have actually read them?

- Is it possible to have an ethical framework that will always fit social media research? Or does it have to be managed on a more case by case basis? When is ethics approval needed? Where should the line be drawn – does social benefit come into it?

- Is ethics committee approval sufficient, since many committees are not sufficiently aware of the ethical issues of social media research or how to deal with them?

- Researchers can influence the shape and structure of data and set the conditions of possibility – what are we actually researching? Is it ethical or not to recognise this influence?
Additional points from Dr Whiteman’s presentation:

- Facebook includes personal information and sensitive topics, it is very complex with respect to public versus private spaces, it links to other organisations; users are often confused about privacy settings and are likely to have an anticipated audience in mind when posting (which is unlikely to include researchers). Other organisations use data that is posted (e.g. law enforcement agencies, journalists). There are issues surrounding the ownership of data. The platform is constantly evolving (e.g. recent introduction of the Timeline feature), making it very difficult to develop lasting conventions of use.

- Research ethics is ‘achieved’ – negotiated, changing – not a set of applied rules. We can’t always feel secure that we’re doing it ‘right’.

- How should what we research be influenced by the researched communities?

- A researcher’s own involvement in social media may inform their approach to using social media data (i.e. think how we would feel if our own data was used in this way).

Key points from the discussion sessions:

The following sets out the four scenarios that were discussed as part of the workshop, and a summary of the key points that were made by the groups that considered them (taken both from the flipchart notes made by the groups themselves, and the notes of specific note-takers recording the outcomes of feedback sessions within larger groups. A full representation of the flipchart notes from each group can be found as an appendix to this document.

Scenario 1: Qualitative ethnography

“The researchers in this project aim to better understand how Facebook is being used by people on the ground in the aftermath of a natural disaster. As part of a wider study, they are planning to conduct an in depth qualitative analysis of a number of Facebook pages that were used by local people to communicate and organise in the aftermath of the disaster. Some of the Facebook pages are private, however anyone can request to join them. Some of the Facebook pages are public. Each private page has at least 1000 Facebook users. There are a wide range of topics being discussed on the boards including people searching for lost family and friends, memorial posts about people who have died in the disaster, information about access to provisions and shelter, and a variety of other conversations between group members. The researchers want to join the private groups, and then observe how different types of public and private Facebook pages are being used by people as they respond to the disaster. They plan to make notes as they are exploring the pages, and take screenshots of any interesting posts and comments. They would like to use these screenshots in future publications and conference presentations.”

Comments made in relation to this scenario:

- It was felt that this was a very sensitive subject involving vulnerable people with significant potential for harm (incorporating personal insights, images and information from private pages, albeit those with a large number of members), and that ethics approval would be required; the research presents more ethical issues than traditional research methods would.
• It was felt that Facebook often blurs the lines between what is public and private, and this lack of distinction is not made at all clear to Facebook users, e.g. the fact that ‘private’ posts can be made public by re-posting.
• Concerns were raised regarding how access to pages and to individuals is negotiated – an ethics application would need to go into detail to clarify how these issues would be handled within the confines relevant tool/platform:
  - Who sets the boundaries?
  - How are the boundaries of access defined?
  - How would researchers negotiate entry/who with?
  - How differently should a researcher treat public and private pages? The users are likely to have little expectation of either being researched.
  - Is it OK for a researcher to join a private page? Should they identify themselves? If there are explicit about doing the research it may distort the use of the group and affect the potential benefits BUT is it ethical to join a group at all without disclosing?
• With at least 1000 Facebook users notifications of research taking place would disappear and users would not be aware that research was taking place.
• If consent was sought does the participant’s vulnerability affect their ability to decide if they want to participate?
• Is this exploiting the vulnerabilities of individuals?
  - Psychological harm
  - Violation of sanctity of privacy
• How could researchers arrive at an ‘ethical base’ – to agree appropriate standards? Could they consult social media users?
• Facebook users are supposed to be at least 13 years old, but it is difficult to ensure that this is the case and many users will be under the age at which they can consent from themselves.
• The project involves very vulnerable groups coping with the aftermath of a disaster, many of whom may be in a highly distressed state; this may result in posting information that they would not normally want to make public. There was also a concern about needing to consider the reputation of the recently deceased, as the pages to be researched serve a memorial function.
• What are the benefits of this research – is it worth it? If research was covert does the benefit outweigh the deception? If the project aimed to help with e.g. better disaster support does this make it justifiable? There may be an obligation to directly support the researched group in some way, or at least future groups in similar situations (but social media changes so quickly – will any recommendations be relevant in the future?)
• If resulting findings/publications are reported widely in the media, there may be a negative impact on the researched community; they may not want this information to resurface, especially images, which can quickly ‘go viral’.
• How would language/translation issues be handled?
• There was a view that this research should be restricted to public pages.
• There was a view that the research could be done in a much less intrusive way. It appears to be taking place to capitalise on an opportunity rather than being done as part of a large well-thought out project = less development time.
• There was a concern about who has the power: the individual or Facebook?
• There was a concern that the use of screenshots was not really necessary - using the text from the posts would mean less chance of identifying people.
• There was a view that the project presents may grey areas, and that in this type of situation the researcher may have to choose the ‘least wrong’ option.
• There was a view that Facebook is a commercial organisation, and data will be used for
commercial research purposes; hence it may be justifiable to do academic research too (but this may depend on risk of harm?)

- There was lack of clarity regarding whether the research involves the researcher engaging with participants or just observing. There may be a need to negotiate relationships/roles beyond as well as during the project.
- In general, there was a view that ethics reviewers of social media research projects should be active social media users who understand social media platforms.

---

Scenario 2: Quantitative analysis of Twitter data using anonymised secondary dataset

"The researchers in this project have been given access to some secondary data by colleagues in another University. The dataset has been derived from a Twitter dataset that was legally generated using an online service, and the original research had ethical approval. The secondary dataset that our researchers have access to has been de-identified - usernames have been replaced with anonymised identifiers, there is no Tweet content other than the hashtags used, and network links (retweets, followers etc) are reported anonymously.

The researchers want to use this dataset to visualise relationships between users of a number of Twitter hashtags identified as having a lot of racist/homophobic/sexist content e.g. #gamergate and #notracist, and users of hashtags of various political campaigns in an election e.g.#whyimvotingukip, #votelabour etc. They aim to identify relationships over time between hashtags they perceive to be racist/homophobic/sexist and the supporters of particular political parties. They plan to produce a number of network visualisations and publish their results in academic publications and the mainstream media."

Comments made in relation to this scenario:

- It was felt that this was a sensitive subject area and hence should require ethics approval.
- There were concerns about reproducibility and validity; there was a feeling that the research was flawed and there were ethical issues attached to the hypothesis of the research that linked membership of a particular political party to racism. The questions seemed incomplete and/or inadequate. Not everyone uses hashtags but may still use racist/sexist/homophobic terms = false selection of data? The researchers should be aware of their own potential biases in undertaking this work. There was a view that the proposed project needed scientific review!
- Although there were some concerns about the topic of the research it was not felt that the research was unduly controversial or in the realms of "wiki leaks". However, there were concerns about whether there were potential benefits to society from doing the study – if so could these justify some aspects of the approach?
- The question of the validity of the dataset was also raised (e.g. under (voting) age participants, humorous/sarcastic responses). There was also a need to examine how and why the hash tag was being used, as there was a risk of use being taken out of context and incorrect assumptions being made. In addition, some hashtags may have been used innocently and the hashtag later morphed into something else (i.e. had a different meaning).
- It was felt that this was a controversial subject which would require anonymization; there were concerns that it might be possible to trace data back to the participants. It was important that users of the # were not victimised and or 'labelled' as racist etc (may have used the # innocently). This raised the question of how responsible are the researchers
for secondary claims? There is a need to protect the following stakeholders

- Participants (some of whom may be children or young people – it was noted that
  things are currently never deleted from Facebook or Twitter and there is a risk
  that something posted when young could affect future life/career)
- The researcher (being associated with extremist material/groups could be a risk;
  could use false profiles to protect identity)
- The institute

- Since the researchers would be receiving a secondary dataset, they may not be aware of
  the sensitive issues until they actually receive the data (i.e. likely to be after ethics
  approval has been obtained).
- The ethics of publication also came under scrutiny; consideration needs to be given to
  what happens when this research is disseminated if the sources are not anonymised. It
  was agreed in one group that whilst there might be a paper published which vanished
  into obscurity there was also the possibility that with such a controversial subject it was
  possible that the media could pick up on the subject and there could be a “witch hunt” of
  followers. The feeling was expressed that the media often override ethical
  considerations with those of “public interest” (the example of Sir Cliff Richard) was cited.
- There were concerns around the legality/permissions required for accessing this data –
  why were the researchers given access – was this open access data?
- There was discussion around informed consent - just because people agreed to join the
  group, does that constitute informed consent? People should be allowed to say no.
- There were concerns about the timing of the research – if undertaken around the time of
  an election then this may present a higher risk.
- There was a view that the ethical approach to research of this type must be developed on
  a case by case basis.
- The scenario prompted a discussion about when something becomes data – e.g. if you
  incorporate your notes/thoughts from a discussion with others into your writing; not
  quotes but what you remember of the discussion – is this data? It was noted that the risk
  of harm needs to be considered.
- A suggestion was made in reference to the UREC guidance to be developed: should the
  UREC consider the ethics applications of social media research that has already been
  approved, and find out if the approaches taken worked or if issues were raised?

**Scenario 3: Content analysis of Twitter data**

“The researchers on this project plan to use an online commercial tool to collect Tweets
using hashtags related to Leicester City Football Club (e.g. #LeicesterCity) over the course of
the 2015/16 football season. The data provider provides its service legally and in line with the
terms and conditions of Twitter. The Twitter data they gather will be fully identifiable. They
plan to manually code the Tweets in order to explore people’s affective (emotional)
response to the club over the course of the season, and how these emotive factors played
out in users’ non-football specific comments (e.g. in relation to general wellbeing,
perceptions of community etc). In the write up of the research, the researchers plan to
include some screenshots of illustrative Tweets.”

**Comments made in relation to this scenario:**
- It was noted that if a researcher wishes to publish Tweets they need to be sure that the
  Tweets haven’t been deleted.
• Is it ethical to use screenshots? There was a concern that individuals and their Tweets should be anonymised to protect individuals, but an analogy was made to letters published in newspapers which would be acceptable to research. Would we protect people who wrote letters to newspapers? Do we need to protect the authors of Tweets?

• It was noted that there may be safety risks to participants e.g. rival fans taking objection to comments made. Also people may not be in a good mental state when posting (e.g. upset at a lost game, under the influence of alcohol/drugs) – would they want these posts to be shared when in a normal state of mind (may have an impact on their future lives/careers)? We wouldn’t normally interview someone when drunk but how would you know if posts were made when under the influence? Could it be seen as eavesdropping on someone in a pub? What would happen if potential criminal activity was identified?

• When people join Twitter are they fully aware of how public their Tweets will be?

• There was a concern that there was no way for people to withdraw from the study.

• There were concerns that judgements would be made around Tweets, which may result in data being misrepresented. The true sentiment may not come across (e.g. sarcasm). The researchers would be presenting their own interpretation; the same may be true of interviewing, but with social media there is unlikely to be an opportunity to clarify with participants what they mean.

• A comparable example was mentioned: companies will sell items which have been produced using third world labour in sweat shops. This is not illegal, it is unethical. Concerns were raised regarding the line between what is legal and what is ethical.

• There was a concern about the brand image of the football club, and what their view would be on this information being collated and published; it might affect their image.

• It was noted that it may be impossible to verify the age of participants.

• If particular individuals were followed over time, it may be possible for the researcher to make assumptions about their state of wellbeing – should concerns be flagged?

• The power relations, and equality and justice were raised as issues.

• In terms of recommendations that an ethics committee could make when approving this project, the possibility of requiring that only statistical information about the posts was discussed, rather than using screenshots of identifiable Tweets (it was noted that there may not really be much value in providing screenshots – research aims can still be achieved and it is not justified by the benefits of the research to publish the actual Tweets.

Scenario 4: Network and sentiment analysis of Twitter data

“The researchers on this project are working in the days after the Charlie Hebdo attacks in Paris and aim to conduct a network analysis and sentiment analysis of Twitter users using the hashtag #jesuischarlie. They plan to use an online commercial tool to collect Tweets. The data provider provides its service legally and in line with the terms and conditions of Twitter. The Twitter data they gather will be fully identifiable.

They plan to create network visualisations that show how particular Tweets became popular through retweeting practices. They also want to visualise how sentiment about the events emerged over time amongst different networks of Twitter users. They want to make an interactive online visualisation in which users will be able to zoom in on particular areas of the network. They plan to embed the original Tweets into the visualisation so that users can explore in depth the specific content of tweets contributing to particular categories of sentiment, and the users who are embedded in particular parts of the network.

Note: Sentiment analysis (also known as opinion mining) refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective
Comments made in relation to this scenario:

- Concern was expressed regarding maintaining the anonymity of Tweeters. This was considered to be a highly sensitive subject and there was clearly a need to protect contributors.
- Concerns were expressed that although Twitter account users signed up for an account, they may not be fully aware that their Tweets were publicly available. This should not constitute informed consent for research to take place on this data.
- There was also concern regarding the bias of the dataset.
- There was a strong feeling that the benefits of this research had to be shown to outweigh the risks.
- Should a researcher engage in social media using their own pre-existing account, or set up a new account to do this? How does the researcher make the role clear if they have pre-existing dialogue or roles?
- Need to clarify in UREC policy: what is personal data? Who are human participants?
- How does the researcher account for non-common or dissenting opinions within a large dataset?
- How does the researcher recognise ‘trolls’ and should they be counted as data or discounted? (recognising that women are more likely to be trolled online)
- How does the researcher recognise uncommon ‘spam’ that looks like a comment?
- Guides/specialists on social media research ethics would be useful, along with a workbook to facilitate the thought process that researchers should go through when developing a social media project.

Comments from Panel Discussion

Overall comments Group A – Helen Kennedy

- The group found it difficult to come to a consensus on the scenarios; they tried to weigh up the benefits and harms but it was difficult to assess the studies fully from the information given.
- They felt that contextual specificity was very important, and that the ethics approval process should be a negotiation; there should be a more dialogic assessment of ethics.
- There should be mechanisms within the institute to decide the level of risk, and guidance on the acceptable level of risk.
- What is considered cutting edge and what goes beyond acceptable? The question is why NOT? What is acceptable may change over time – ethics is societal; ethics reflects society (research which took place 50 years ago would not be acceptable now).
- Anonymisation of data was a key issue, and the use of screenshots – it was noted that individuals may become identifiable from hashtags where there is a small group of users.
- Although data may be available publicly, if researchers would ultimately be re-publishing the data, there should be ethical consideration.
- Concern was raised regarding what is done with data that is generated.
- It is important that researchers DO NOT set the boundaries of what is ethical.
- Research should gain consensus.
- Key questions to ask - are we protecting:
  - The source
  - The researcher
  - The individual
**Overall comments Group B – Jo Bates**

- Key issues discussed were around the public/private divide, anonymisation and consent.
- The group felt that scenarios 1 and 4 would definitely need ethics approval.
- Even where ethics approval is obtained, there is a need to think continually about what is ethical as you go through to publication.
- Ethics for social media research – involves more questions than answers, complexity. There a need for a more discursive process around ethics in the process itself, and in training.
- Strict guidelines would not be workable – a case by case approach to ethics is better.

**Overall comments Group C – Peter Bath**

- There is a need to consider the potential harm to individuals or groups (especially well-being and safety of vulnerable groups e.g. those in extreme circumstances).
- Public vs Private – ‘rules’ need to be applied in a different way.
- If there is a potential benefit for those being researched, does this mitigate some risks?
- Is it more acceptable to research public areas of social media?
- Overlap between legal (e.g. data protection) and ethical issues - what is legally acceptable is not necessarily ethically acceptable.
- Consideration needs to be given to the potential consequences of actions.
- Would we be happy for this to happen to our own data?

**Natasha Whiteman – summary of key points/thoughts**

- People need a greater awareness of context that what they Tweet/post will go out into the wider world – can the research community help with this?
- Consent is a key issue
- Public versus private is a key issue
- Temporality needs consideration: what happens to the relationships after the research has finished (e.g. if a researcher develops a relationship with someone for research purposes – e.g. ‘Friends’ someone); deleted Tweets - how will researcher know this?)
- Policy should set expectations for applicants and reviewers – may need a more rigorous review system
- Disclosure is a key issue (should the researcher disclose their identity/purpose?)
- It’s about protection for the researcher/institution – not being sued – reputational risk

**Closing questions from attendees**

- Should all internet research be classified as high risk as suggested by ESRC’s Framework for Research Ethics (may be considered covert, may involve children)? (Panel response: hesitant to label all in that way; requires reflection on a case by case basis).
- Staff need to be equipped to be able to deal with what we are presented with e.g. supervisors of student projects may not be aware of these issues. UREC and ethics coordinators in departments should expect students/supervisors NOT to know about this area and proactively communicate/ build into teaching – at least sufficient to recognise there are issues and to know where to seek advice (UREC response: All should have read the Ethics Policy – if concerned, the supervisor should flag as high risk so someone else reviews as well (for UG/PGT) or for PGR 3 people will review = layers of
scrutiny should reduce the risks. May need to adjust the risk indicators in the ethics forms).

• No review/audit is currently required after ethics approval is granted, only the data collection and method is considered, but what is analysed and presented need to be addressed as this is what influences policy and practice. (UREC response: this issue relates more to research integrity more broadly – the University has a separate policy on this).

Lindsay Unwin & Anita Kenny
Research & Innovation Services
August 2016
Appendix 1: Flipchart notes made by groups

Scenario 1: Qualitative ethnography

- Lots of ethical issues, so needs review and approval. Clearly involves participants so ethics approval necessary
- Involves vulnerable participants put high risk
- Questions about benefits of this research – who benefits?
- Would need detailed and well defined research questions
- Potential harms
  - Screenshots violate privacy
  - Limits benefits of the group
- Confidentiality / consent
  - Negotiating access
  - Issues over access to private sites and distorting content
  - Covert/overt
  - Revealing personal and sensitive information
  - Not ethical to screen grab pages from private groups
  - Yes, would be happy for public social media to be used in the data set
  - Guidance needed on how to protect
    - Under 18s
    - Reputation of the deceased
- Key issue:
  - Public vs private spaces
    - Users added all the time
    - Statement of intent might not be seen

Consent

- ‘screenshots’: these can change
- Vulnerable participants, signposting to support
- Consent from deceased participants
- Confidentiality, records of conversations
- Researcher declares their role?
- Did they have previous role in the group?
- The moderator gives consent on behalf of ALL
- Moderator can’t contact all individually
- Future joiners don’t consent
- Are the incentives too high to participate/help?
- Participant assumptions re help researcher can provide
Appendix 1: Flipchart notes made by groups

Scenario 2 Quantitative analysis of Twitter data using anonymised secondary dataset

1. Requires ethical approval
   (Why not?)
   – Yes, high risk area of research
   – Better protection
   – Peer scrutiny and dialogue

   Description is incomplete / inadequate
   Consider ethics of data collection, aims, publication (e.g. claims about impact) risks to subjects and researchers.

2. Hashtags could put users at risk
   – Linking hashtags to small user groups
   – Anonymise hashtags – e.g. “Hashtag A”

Ethics of access to social media

Ethics of experimental design

Ethics of T&C vs. people’s understanding

Ethics of children using social media & anonymisation / ignoring age

• Ethical approval needed? UREC
  – Yes, high risk area of research
  – Purpose of original research may have been different / new considerations
  – Data set modified? # out of context What data cleaning has taken place?

• Potential harm
  – Difficult to assess – depends on size of dataset. Small dataset = easier to reverse engineer & identify users.
  – Users may not be aware of negative connotations of # - they are then labelled as ‘racist’
  – Young may be particularly vulnerable – No demographic information in the dataset
  – May badge them negatively
  – What has defined these # as ‘racist’?

• Confidentiality and Consent
  – Would need to go back to source data?
  – Find out what the user consented to at the time of data capture? Terms and conditions.
  – Is the data ‘open access’ – why is it being passed on? Need to be clear about this
  – Context of tweets is important / tweets may be deleted. Need to address this.
  – Is the data from a ‘real’ identity?

• Consider is informed consent really ‘informed’ consent?
• Risk vs benefits?
• Would you be happy?
  – No – those using twitter socially/humorously / without consideration of the implications
    – may be damaging / misleading
  – Inferred affiliations
Appendix 1: Flipchart notes made by groups

- Freedom of speech vs privacy
- Social media – like ‘chatting in the pub, not a ‘serious’ or deliberate forum.

- Guidance from the UREC
  - Secondary use of data
  - Guidelines for individuals accessing data (secondary data rules)

- Sensitive research – high risk areas need much tighter guidance as impact greater
- Context – when the data was captured / needs defining in the context e.g. Post Brexit
- Risk and Impact assessment, what is the purpose / impact of disseminating the results?
- Review first social media approvals for issues and best practice
Appendix 1: Flipchart notes made by groups

Scenario 3: Content analysis of Twitter data

- Researcher knowledge may be more useful in this situation (More specific)
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
- Unclear
- May depend on how tweets are used e.g. deleted tweets?
- Can take a long time
- Using scenarios would be helpful
- Ethical approval?
- Possibly not because of platform terms
- “no one forces you to tweet about breakfast
- It’s about self-publication It wouldn’t work if it was private
- Research can be dubious but largely depends on data, how it’s used and expectations
- More general framework so can be grey area
- UREC guidance
Appendix 1: Flipchart notes made by groups

Scenario 3: Content analysis of Twitter data

1. Ethical approval
   - Yes! Institutional / ESRC
   - How valuable is the research in relation to risk?

2. Unintentional Consequences
   - Leicester City – Brand
   - New set of meta data
   - Level scrutiny that is not expected.
   - New audience

3. Identifiable
   - Dangerous (from opposing fans) use pseudonyms
   - Press involvement
   - Right to withdraw
   - Tweets taken out of context
   - Validity of the data
   - Group accounts
   - Criminal implications
   - Covert observation
   - Data from vulnerable people – drugs, alcohol / trauma?

• Methodology
  - Gathering data okay but how it is published could be a condition
  - Why do you have to screen shot the tweets
  - Could stats be better
Appendix 1: Flipchart notes made by groups

Scenario 4: Network and sentiment analysis of Twitter data

- The researcher should obtain ethical approval
- Confidentiality – the tweeter identified should be anonymised to protect that person from potential harm
- Does the benefits of the research outweigh the risks?
  [risks identified relating to all stakeholder groups]
- Consent – not practical:
  - Too many “participants”
  - Informed consent – limited understanding is an issue
  - Expectations
- Unforeseen consequences:
  - Misunderstandings – the problems of language and a global platform
  - Cultural perspectives: meaning may not be as intended by the participant / interpreted by the researcher
  - Learning by machine? Inaccuracies compounded potentially. Sentiment may be misinterpreted
- Twitter data: use mine?
  - Yes in a professional capacity
  - Maybe in a personal context

Consent: key issue, amalgamated data vs identifiable data

  - Contexts, possible /impossible of the practical

Methods: knowledge of what is being done and how.

- Twitter privacy – public vs private
  - What counts? Where do they intersect/blur? Subjectivity of user?
- Practice of using hashtags: subjective/individual
- Paraphrasing – methodological issues – sentiment analysis – ethically, does paraphrasing remove/alter the sentiment?
- Deleted tweets – post scraping. Does the researcher go back and find (and delete) tweets?
- How is the analysis processed, and turned into a visualisation?
- Analytical tool (see data visualisation) drawing attention to certain people / emotions, especially because it’s a controversial issue.
- What is the value of the research?
- Personal data – what counts?
  - Ownership of tweets.