



Usage Protocols and Guidance

The legal and ethical aspects of using online platforms, inc social media, when undertaking research under the UOL auspices

Contents

Introduction	2
Online Research Methods: Definition and Classification	2
Online Research Ethics: Advice per methodology	2
1. Observational or Non-intrusive Web-Based Research.....	3
2. Interactive Research Involving Engagement with Participants	3
3. Surveys and Interviews	4
4. Collecting and keeping data from social media sources.....	5
Advice per platform:	7
1. Social Networks (eg: Facebook, LinkedIn, Twitter, Instagram).....	7
2. Media Sharing (eg: Pinterest, YouTube, Vimeo).....	11
3. Blog comments and forums	11
4. Communication tools (Zoom, Skype, Whatsapp, Blue Jeans, Microsoft Team Meetings...)	12
Key areas of concern within social media research	13
1. Private vs. public?	13
2. Informed consent.....	14
3. Anonymity	14
4. Risk of harm	15
5. Terms, conditions and legalities.....	15
6. Privacy and risk	16
7. Re-use and re-publication.....	18
8. Are the data collected protected by IPR?	18
Summary	20
Further reading and other sources	20

Introduction

1. It is recognised that social media provides numerous benefits and opportunities for researchers to use available platforms.
2. However, along with these benefits come the risks of managing something that is dynamic and potentially unlimited in scale. These include the reputational damage arising from misuse, abuse, threats to the security of sensitive or confidential information, exposure to malware and disruption.
3. This guidance is designed to support all researchers (staff, fellows, students) to use online platforms, inc social media appropriately and to protect their rights as researchers as well as those of their participants, as well as the reputation of the University.

Online Research Methods: Definition and Classification

In the past few years, many researchers and organizations have attempted to define and classify online or internet-based research. According to the Association of Internet Researchers, internet research encompasses enquiry that:¹

1. utilizes the internet to collect data or information, e.g., through online interviews, surveys, archiving, or automated means of data scraping;
2. studies how people use and access the internet, e.g., through collecting and observing activities or participating on social network sites, web sites, blogs, games, virtual worlds, or other online environments or contexts;
3. utilizes or engages in data processing, analysis, or storage of datasets, databanks, and/or repositories available via the internet;
4. studies software, code, and internet technologies;
5. examines the design or structures of systems, interfaces, pages, and elements;
6. employs visual and textual analysis, semiotic analysis, content analysis, or other methods of analysis to study the web and/or internet-facilitated images, writings, and media forms;
7. studies large scale production, use, and regulation of the internet by governments, industries, corporations, and military forces.

Online Research Ethics: Advice per methodology

The general principles that guide ethical practice in online research are essentially the same as those that guide any research involving human beings.

- Each individual has the right to privacy and dignity that should be protected at all times. As such, every participant should be able to make their own decisions to participate in research and the persons who are unable to make these decisions should be protected. Even in the context of online research, researchers must protect the personal information of internet users and refrain from disclosing anything (intentionally or not) that would allow their personal information to be accessible. As such, participation sheet and consent form must still be used.
- All research participants should be treated fairly, equally, and nobly during the entire course of the research process. This means, that the researcher's identity and the

¹ (Markham & Buchanan, 2012, p. 3f)

research methods must be transparent, and that no-one contributing to the research is unfairly treated or faces discrimination. It also means that the researcher should protect all participants' own interests (in particular if they cannot do it themselves), including protecting from any exploitation for the sake of research (World Medical Association, 2006). As such, online recruitment should be truthful, transparent, and honest when describing the aims, details, risks, and benefits of their studies.

- Researchers must evaluate all physical, social, psychological or medical harms or risks that their participants may face by virtue of being in the project, and make every possible attempt to minimize these harms and maximize the benefits to them. Within the context of online research, the risk of harm arises when there is a disclosure of participant's identity or any other sensitive information that may expose them to the risk of embarrassment, reputational damage, or legal prosecution. (which can happen by choosing a particular online platform)

Overall, each type of online research method (i.e., observational, interactive, or survey/interview research) is highly contextual and involves different levels of engagement and interaction between the participant and the researcher, which has implications for ethics. **Our recommendation is to apply risk mitigation strategies that are proportional to the magnitude and probability of risks, which are clearly outlined to the UOL Research Ethics Committee.**

1. Observational or Non-intrusive Web-Based Research

In this type of online research, participants are not aware of the research, except when information obtained is retrieved or recorded in such a manner, which identifies the people, directly or indirectly, and/or puts them at risk of civic liability, or be damaging to their reputation. This particular category applies when investigator observe social websites such as Facebook or Twitter, access public blogs, or download and/or analyze YouTube videos.

Some elements that are of particular importance for the researchers undertaking an online observational study are users' privacy, the website privacy policy, and other legal considerations – see below for further details. The users of social media platforms have the choice to make their posted content private or public, which in turn becomes the deciding factor whether the study is exempted from ethics consideration or otherwise.

In situations where content is public (i.e., does not need password or permissions for access), **the researcher must determine whether the use of content** (e.g., quotations from blogs) **reveal the identity of the author/owner of the content.** If this is the case, the researcher **must seek informed consent, minimize the use of direct quotes/content, and consider a certain level of disguise.** Most of the social websites such as Twitter, Facebook, and LinkedIn have their privacy policy. If researchers intend to use these websites, they should either certify compliance with the terms of use on the sites or alert the institutional ethics review board whether the techniques comply with the terms of use or not. Some of the legal considerations relevant to online observational studies relate to the question of what constitutes a privacy violation.

2. Interactive Research Involving Engagement with Participants

In this type of online research, researchers identify themselves and directly contact the participants, with an intent to access their online media content (such as social interactions) that is not publicly available, or interact with a person (e.g., through friending or following), or gain membership in a closed group (such as a chat room or Facebook group). Some authors refer this type of research as virtual or visible ethnography. It is recommended to make direct communication with the participant. **The researcher must ensure anonymity of the**

subjects and gain informed consent without any coercion, as in normal circumstances, to protect the privacy of their participants, but also the privacy of the website/media (such as chatrooms or blogs) through which they draw their information/data. If participants are viewed as authors, the researcher may decide to give acknowledgment where it is due. If however, the participants are regarded as “subjects,” the researcher needs to guarantee their anonymity.

3. Surveys and Interviews

Surveys and interviews are common and well-developed methods of online research. In conducting surveys and interviews online, ethical concerns may arise with regard to obtaining informed consent, ensuring anonymity of the participants, and maintaining confidentiality of the data. Consent is considered to be knowledgeable when individuals know the purpose of the information collected about them and their right to give, withhold, or withdraw the consent anytime they wish.

Enrolling participants for research involves some open and clear communication between the researcher and the potential participants, however this may be compromised in virtual settings. Different ways through which an online consent can be obtained include emails or online statements that may require participants to agree to the terms of participation, thus implying consent. However, it is difficult to verify whether the participant has actually read the details carefully, whether there are any misunderstandings and issues in comprehension, and whether the person who is giving the consent is the authentic participant. Furthermore, verifying certain information such as age or mental capacity to give the consent can also be difficult.

Researchers should provide complete details of the study **as in normal circumstances** (see Ethics guidelines for details), including a readily available link that provides contact information, study aims, data collection procedure, potential benefits and harms, and steps taken to maintain anonymity and confidentiality of the participants.

A few ways to ensure participants read and understand the study details in online settings include use of multistage consent forms, enhancing readability of the document by reducing the amount of text, use of subheadings, or colours (Social Sciences and Humanities Research Ethics Special Working Committee [SSHWC], 2008). Testing the consent forms and survey questionnaires are other important ways to minimize the risk of unintended consequences at later stages of the research.

Preserving Anonymity in Online Survey Research

It is the prime responsibility of the researcher to ensure anonymity and confidentiality of the data, which is collected and stored. If the data are not contentious, or if privacy can be ensured, then this is less of a concern compared to controversial research topics or research where it is necessary to obtain personal information. In case of any sensitive information, researchers need to ensure that participants’ perceptions of anonymity are met, or if not, made explicit to the participant. Steps to take to ensure ethical practice include:

1. Survey respondents’ participation must be voluntary at all times.
2. Personal information should not be sought from or about the respondents without their prior knowledge.
3. The researcher must ensure that data are not used for subsequent non-research purposes.
4. Personally identifiable data and other research related data should be kept and stored separately.

5. Respondents should be informed about the duration of the survey and must be given links to data protection, privacy policy, cookie policy statements, or information about the use of a software at the beginning of the survey, which should be available in simple language.
6. Any software used for the survey should not be installed on the respondent's computer without their consent and even when consent is provided, respondents should be able to remove any such software easily.
7. Respondents should be given entitlement to ask the researcher to delete their records and researcher should conform to such requests where reasonable.
8. In case of surveys where an email list is acquired from some agency, the researcher should provide a clear statement of where the email addresses came from and ensure that the individuals listed have consented before to be contacted in future for any research purposes.

Data Protection in Online Survey Research

Any researcher should aim to abide by appropriate principles under which any personal information should be collected and secured, including physical, and technical safeguards to protect the personal information against theft, loss, and unauthorized collection, use, copying, modification, or disclosure. Some of the principles include the following:

1. Personal data collected or stored on the websites or servers should be protected using measures such as data encryption or Secure Socket Layer (SSL).
2. For ensuring physical protection of the data, researchers should keep the data in password protected computer directories, use data labels that are meaningless to anyone but the researcher, and code the data in a way that reduces the possibility of people being able to trace the data to a specific individual.
3. In case of temporary storage of data on the server, the researcher is obliged to take necessary precautions to prevent unauthorized access to the data from the server or during data transfer, and ensure that the temporary storage is terminated at the earliest.
4. In their privacy policy or statement, researchers should always state the use of cookies or other log files and inform the respondent that they can disable use of cookies in their computer by using their own system settings.
5. When data are collected through the server that does not belong to UOL, researchers must ensure that compliance to UK and EU privacy laws can still be ensured.
6. Researchers must observe all relevant laws and national codes, specifically if their projects involve participants from various parts of the world.
7. They also need to ensure that data transfer is permissible and is safeguarded against any breaches to privacy.
8. In case where emails are sent in batches for the purpose of recruitment or inviting the participants for surveys, researchers need to ensure that the emails of the respondents are not revealed. If possible, the use of emails should be minimized.
9. In case of any breach to data safety or anonymity, the UOL Data Protection Officer should be informed immediately with all the details. Based on the circumstances of the incident may be necessary to notify the participants.

4. Collecting and keeping data from social media sources

The harvesting of social media data can qualify as “processing personal data”. In addition to processing names, social media platforms will often include text, videos, audio and images

etc. that can be entered into a search engine and traced back to a living individual, and will thus be subject to the General Data Protection Regulation (GDPR). If data contains personal information, make sure to be compliant with the privacy rules.

Personal data exchanged in public messages on social media platforms isn't owned by the brand or the agencies acting on behalf of brands; it is owned by the individual who uses social media platforms. The social media platforms have their own privacy notices and guidelines which social media platform users and advertisers agree to comply with. On top of this, brands must set out in their privacy notices how they will use such personal data in accordance with the requirements of the GDPR, in particular, but not exclusively, the right to be informed. An example of this can be the Information Commissioner's Office's social media section in its privacy notice.

The GDPR only applies to individuals' personal data and not to any information about organisations. However, contact details of organisations' members (employees, contractors, consultants, trainees etc) such as name, job title, phone number, email address, and personal social media account details all fall within the definition of personal data.

GDPR legislation stipulates that personal data must be collected for "specified, explicit and legitimate purposes". Therefore, when personal data is collected, website owners must first explain to visitors how it will be used and provide them with their information rights, and secondly ensure the different purposes for processing the personal data are separated out.

It is the researcher' responsibility to specify in what way the personal data will be used and ensure it is not ambiguous.

The GDPR lists a number of rights to individuals in regards to their data. There are, however, a number of exemptions to these rights in regards to research data in the GDPR and the UK Data Protection Act 2018, as long as appropriate safeguards are in place. Details can be found in Section 14 Appendix C of our [GDPR guidance](#).

The processing of personal data that takes place on a social media platform is ultimately the responsibility of the platform provider. If the user agrees to these terms and conditions of use, and the data privacy notice, then it's up to that platform owner to ensure that the personal data held on the platform is processed in compliance with the GDPR. However, by removing personal data from the original platform on which it was published, the responsibility shifts to those who become data controllers (the researchers). **Accordingly, we have to treat that personal data in the same ways as any other personal data under the GDPR.**

Ultimately you will have to comply with all the legal requirements under the GDPR including but not limited to making sure you have a legal ground under the GDPR for processing such personal data; a valid purpose as to why you are retrieving the personal data from the social media platform; and being able to justify your actions under the accountability principle. The researcher's legal ground is

For the performance of a task in the public interest or in the exercise of official authority vested in the controller

See our policy [here](#)

If a researcher is storing personal data, he/she needs to:

- ensure it is accurate and kept up to date.

- only keep the personal data for as long as it is necessary for the specified purposes.

If you export a spreadsheet with all the posts on it from a social media platform on it (highly likely to contain personal info), ask how you'll store it: do you really need it at all

For further information regarding academic research and data protection, please review our extensive guidance here: <https://www.sas.ac.uk/research/research-policies-and-protocols/research-ethics>

Advice per platform:

1. Social Networks (eg: Facebook, LinkedIn, Twitter, Instagram)

A social networking site is a social media site that allows you to connect with people who have similar interests and backgrounds. These platforms allow you to connect with friends, family, and even brands, as well as organisations. Most social network sites let users share thoughts, upload photos and videos, and participate in groups of interest. **They all have terms and conditions of usage, which need to be understood before using the platform as part of academic related activities undertaken under the University's auspices.**

Facebook

Unfortunately, Facebook offers participants a relatively high degree of control over their data, **but it is the researcher's responsibility to weigh the costs and benefits of collecting and using personal user information — and to defer to an ethical review when in doubt. The mere availability of data and participants' willingness to share them does not grant researchers the right to record and use them freely.**

The lack of formal guidelines is exacerbated by ever-accelerating technological progress; both researchers and Research Ethics Committee members may over- or underestimate the threats to participants, thereby hindering benign projects or approving malignant ones. Both factors discourage social scientists from conducting online research or submitting studies for review.

There are two major ethical challenges pertaining to data collection in the Facebook environment.

- First, the boundary between data belonging solely to participants and information belonging to others is very vague. Participants' consent allows researchers to record content that refers to or was contributed by other people, such as tagged pictures, videos, messages, or comments on the participant's profile. In our view, it is acceptable to use data generated by or containing references to nonparticipants, but only if the analyses are aimed exclusively at those directly participating in the study. For example, nonparticipants' demographic profiles and network connections could be used to establish the parameters of a participant's egocentric social networks, or the gender ratio among their friends, but not to study any of the nonparticipating friends.
- The second major challenge refers to the vague boundary between public and private information. Some basic profile information is publicly available and even indexed by search engines. However, some scholars point out that the border

between public and private is not determined by accessibility, but by social norms and practices. For instance, in a small town where everyone knows intimate details about everyone else, people tend to pretend not to know facts that are considered personal. Others argue that mining public data is equivalent to conducting archival research, a method frequently employed in disciplines, such as history, art criticism and literature, which rarely involve rules for the protection of human subjects.

Our advice is that public Facebook profile data may be used **without participants' consent if it is reasonable to assume that the data were knowingly made public by the individuals. Researchers should, however, immediately and irreversibly anonymize the data and abstain from any communication or interaction with the individuals in the sample. Furthermore, researchers should be cautious not to reveal any information that could be attributed to a single individual (such as photographs or samples of text) while publishing the results of the study.**

We recommend that you:

- **Create a Facebook research page**

It's a good idea to create a Facebook page for your research group or organisation, to keep a professional presence on Facebook that is **separate** from your personal one.

People can withdraw at any time. Information and consent forms are always available.

- **Polling**

For simple research, just click the "Questions" tab and then write the question you want to ask. You can leave the answer open or add options they can select from. For more options you can use the Facebook Poll app which can be found on Facebook, using the Search bar.

- **Events**

One really useful feature in Facebook is the Events feature. People are increasingly using this tool to organise and publicise informal events and even conferences on Facebook.

When you create the event on Facebook, you have three choices:

- Open – Events can be seen by anyone, and anyone can send invitations to the event.
- Closed – The event listing is visible to anyone but only those invited can see the details.
- Secret – These events can be seen only by those who are invited.

There are other options as well. If you want to make the event more interactive, you can enable the 'Event wall' and allow other people to post pictures, videos and links. This is useful to get feedback and to keep in touch even after the event has ended. Go through the [Events Help Centre](#) for more information on Facebook events.

- **Groups**

To communicate with people who are not already your Page's Facebook friends/followers, you can recruit respondents via Facebook groups. These allow you to talk to members in real time or via email. Updates will come to your email like a regular mailing list. You can post an invitation to the group along with a link to your survey site.

People can withdraw at any time. Information and consent forms are always available.

Facebook can be used to remain in contact with ex-participants. Their comments provided the authors with invaluable feedback on the design of and issues with the studies.

- **Questionnaire**

An existing online survey or questionnaire can be integrated easily with Facebook by adding a fragment of HTML code. Obtaining access to participants' Facebook profiles means that many of the typical questions (such as those concerning demographics) can be skipped given that data can be obtained directly from the Facebook profile or inferred from the targeting approach used to promote the link to the study.

- **Incentives or not?**

As in other contexts, offering appropriate incentives to participants is an important consideration while designing a Facebook-based study. In general, we discourage financial incentives, as they do not reward people for responding honestly or behaving naturally, but merely for participating in the study. Thus, not only are financial incentives expensive, but they may also encourage dishonest or random responding, and attract semi-professional participants. Rewarding participants with an enjoyable experience or interesting feedback can achieve a much better alignment of the participants' and researchers' interests. Even the studies that do not produce interesting feedback could be enriched with elements that do.

- **Useful tips:**

- Think of designing the study around the needs of the participants' experience. For example, participants are often prevented from skipping questions or tasks, or barred from accessing the study based on their demographics. Such an approach may be likely to trigger dishonest responses or behaviors. In practice, it is usually easier to remove participants and protocols that do not meet the criteria before the analysis, rather than trying to filter them out at the data-collection stage.
- Some people might not be comfortable with being used as research subjects, so reassure them that your research is for academic purposes only and that responses will be anonymous.
- Credibility can be an issue: if you have a fake name with a cartoon as profile picture, it might be difficult to get responses.
- Because Facebook participants come from diverse backgrounds, they may misunderstand instructions or test questions due to linguistic or cultural differences.
- Please be mindful of the profile you create as part of the research. You may wish to separate your personal profile to your work profile. Facebook profile information includes self-reported information and data contributed by others (such as photo tags or comments on a user's wall). It is recommended that the research page be attached to a profile on a UOL/Institute website to confirm that the researcher is real.

The lack of face-to-face contact increases the psychological distance between the researcher(s) and participants, which may decrease the participants' feeling of accountability, which would be wrong.

Papers employing Facebook data should include a discussion of ethical considerations related to the design of a study and its findings. Such an approach ensures that the authors have considered the ethical aspects of their own work, and supports the evolution of standards and norms in this quickly changing technological environment.

Twitter

Twitter is best described as a cross between blogging (or micro-blogging) and instant messaging. It allows the user to send a short message of not more than 140 characters (known as a tweet). Whereas conventional blogging is often confined to a single author, Twitter is more interactive: it can create a conversation that anyone can join.

One of its fundamental principles is that you elect to follow those whose conversations most interest you: these become your "Twitter feed". Twitter's immediacy has meant that it has played a key role in many contemporary events, which is of interest to the academic community in many ways.

Some advice on creating an account is available in Part 1 (of Social Media guidance).

Twitter is excellent for obtaining data from a survey: you can set up a micro-blogging site and get responses from thousands of participants.

Twitter creates an automatic database of information in real time, which is archived and has become a unique source of historical information.

Researchers must be clear on their methodology as to whether they collect information that is on the platform or guide the discussion by for example using hashtags. Either way, this approach would resemble 'covert observation' methodology, whereby the participants are unknowingly taking part and their comments are used. The academic purpose and the terms and conditions of Twitter may allow the researchers to use the tweets in this way. Indeed, by agreeing to the terms, users will consent for their information to be collected and used by third parties (Twitter, 2016A and B). A justification often provided with regard to the ethical and legal implications of using data without informed consent is that the reuse of data is permitted in the Terms and service as well as within the privacy policy.

But several aspects need to be considered, as mentioned above, inc.:

- The boundaries of what is private vs public comments;
- The lack of interactions with knowing participants;
- The fact that all comments can be deleted even if archived;
- The lack of anonymity when tweets and comments are being used outside Twitter;
- The participation pool, which is large, and varied, and can include vulnerable people....

It is also important to note that the act of scraping downloading tweets from Twitter's advance search will contravene Twitter's terms and conditions, therefore voiding any protection these policies are likely to offer. This procedure would bypass retrieving data from Twitter's APUS and would allow Twitter to see who has retrieved data from the platform. As a consequence this practice is expressly discouraged by Twitter: scraping the Services without the prior consent of Twitter is expressly prohibited (Twitter, 2016B).

Additionally, reproducing tweets but removing user IDs, or altering tweets significantly will contravene Twitter's user Development Policy, which requires tweets to be published in full. Policies are available here:

- <https://developer.twitter.com/en/use-cases/academic-researchers>
- <https://developer.twitter.com/en/developer-terms/policy>

Twitter users may be concerned about of who owns tweets and whether users have the right of ownership and copyright of tweets that they post. However, in practice, many tweets would not be considered under copyright law because, in most cases, tweets are not original messages (Shinen Law Corporation, n.d.). Twitter's policy on **Content redistribution**:

- ***The best place to get Twitter Content is directly from Twitter. Consequently, we restrict the redistribution of Twitter Content to third parties.*** *If you provide Twitter Content to third parties, including downloadable datasets or via an API, you may only distribute Tweet IDs, Direct Message IDs, and/or User IDs (except as described below). We also grant special permissions to academic researchers sharing Tweet IDs and User IDs for non-commercial research purposes.*
- *Academic researchers are permitted to distribute an unlimited number of Tweet IDs and/or User IDs if they are doing so on behalf of an academic institution and for the sole purpose of non-commercial research. For example, you are permitted to share an unlimited number of Tweet IDs for the purpose of enabling peer review or validation of your research. If you have questions about whether your use case qualifies under this category please submit a request via the [API Policy Support form](#).*

However, researchers need to act more ethically than other investigators for professional reasons.

2. Media Sharing (eg: Pinterest, YouTube, Vimeo)

Media sharing websites allow users to share different types of media, with the two main ones being image sharing and video hosting sites.

Most of these sites also offer social features, like the ability to create profiles and the option of commenting on the uploaded images or videos. These platforms mostly encourage user-generated content where anyone can create, curate, and share the creativity that speaks about them or spark conversations.

If the media platform is used in the context of research, we recommend, following the same advice as the one outlined for Facebook. This is therefore easier to manage participation and participants' expectations within an ethically sound methodology. It is the responsibility of the researcher to ensure that they have read the terms and conditions of the platform before undertaking their research.

3. Blog comments and forums

An online forum is a site that lets users engage in conversations by posting and responding to community messages. The comments are usually centred around the specific subject of the attached blog.

Researchers can set up within the infrastructure of the University of London, to manage their research and lead questionnaires or surveys.

Reminder: Online questionnaires and surveys must follow UOL guidance and include:

- information on the research
- consent form
- ethics approval reference number

- contact details of researcher
- contact details of the Research Office
- Privacy notice.

4. Communication tools (Zoom, Skype, Whatsapp, Blue Jeans, Microsoft Team Meetings...)

There are ethical and practical challenges of online research interviews

Researchers routinely use Skype and other software to conduct interviews – especially when there are challenges accessing the field site. However, this does not stop the researchers to ensure that all aspects of their methodology are ethically sound, in particular as they can still:

- provide information on the research with ethics approval reference number, contact details of researcher, contact details of the Research Office and Privacy notice.
- And ask participant to sign consent form

The flexibility of online interviewing is very clear. However, the physical separation of researcher and participant, and if needed the interpreter is one of the main practical and methodological challenges of online interviewing.

In addition, some of the terms and conditions of using of such tools because of their potential extensive rights to hold and publish our content as they wish. For example, whilst researchers will always own their content, Zoom do not bind their publishing rights to the delivery of the service. Therefore, legally they can re-publish anything put through their services anywhere they choose to. This combined with the extensive mining of users' personal devices when connecting to the service, makes it difficult to sustain, when thinking about anything that is not already in the public domain. As such, the University strongly recommends **against the use of Zoom when undertaking university business, including research**. Should researchers wish to use such a platform or any other platform with the same policies, **they need to agree that their personal devices as well as their participants would be subject to Zoom's data mining, and as such transparency of methodology and consent from participants must be secured prior to the meeting taking place**.

High-speed Internet access and computer literacy of all parties are not always a given. **Potential technical challenges** included sound quality or webcam issues, a time-lag in the audio/video feed meaning sound and/or video is relayed slower than real time, and potentially lost data as a result of technological failure. It **is recommended to try and test before undertaking the interview**, including in a group setting. The researcher can label the session as the **informed consent session** to provide all parties with an introduction to the online format, as well as providing participants with an introduction to the mediation of conversations by an interpreter or gatekeeper (if used).

Conducting interviews online means adapting researchers' ethical obligations towards their participants. They can no longer fully guarantee the confidentiality of information shared and the privacy of conversations, as they often have no control over the location from which participants conduct interviews, or over the way internet platforms retain information. Yet, establishing a "safe" online environment can act to encourage participant disclosure in interviews. To help achieve this safety, it is recommended to create an online secure, password-protected link.

Another issue is **securing physical space and privacy for participants**. The online format means the researcher cannot always control the participant's environment to ensure confidentiality. For instance, researchers should advise participants to avoid a shared office

or public space, which can lead to interruptions or the presence of others in the background. But it is recognised that matching space with accessible and reliable internet access may be difficult for participants, and those may need a gatekeeper to support them.

The lack of physical presence during interviews could lead to misinterpret visual cues, such as smiles or turning of heads, which could be non-verbal cues relating to the conversation, or a response to the presence of others in the room. To meet this challenge, it is encouraged for participants to use earphones, but it is possible participants self-censored their responses for fear of saying the “wrong” thing. To try to mitigate this, it is recommended to keep notes about perceptions of what is happening in the environment around the participant to document the potential role of the local surroundings.

Aside from the lack of in-person interaction in online interviews, the presence of an interpreter, an additional unknown third party, may also have impacted upon rapport building. The use of mainly male interpreters whilst the majority of participants interviewed online were female may also potentially have influenced narratives.

Online interviewing is not a flawless technique to research data and to protect participants. Researchers cannot protect participants in the same way as during a physical interview, as they are not in the same position of control. Online interviewing should be used as an alternative when the physical interview is impossible.

Key areas of concern within social media research

1. Private vs. public?

To assess if the source is public or private, such things as requirements to register, the presence of a moderator, password protection, and sensitivity of content etc. would all suggest some intent toward private communication. Open access, institutional accounts, and broadcast messages all suggest more public intentions. Whether the users generating social media content consider their posts to be public or private will be one of the factors determining whether informed consent for use of the content is required.

One of the biggest areas of concern with social media data is the extent to whether such data should be considered public or private data. Key to this argument is the standpoint that social media users have all agreed to a set of terms and conditions for each social media platform that they use, and within these terms and conditions there are often contained clauses on how one’s data may be accessed by third parties, including researchers. Surely, if users have agreed to these terms, the data can be considered in the public domain?

The process of evaluating the research ethics cannot be ignored simply because the data are seemingly public. Questions of whether online postings are public or private are determined to some extent by the online setting itself, and whether there is a reasonable expectation of privacy on behalf of the social media user (British Psychological Society 2013) – for example a password protected ‘private’ Facebook group can be considered private, whereas an open discussion on Twitter in which people broadcast their opinions using a hashtag (in order to associate their thoughts on a subject with others’ thoughts on the same subject) can be considered public.

Questions of whether the data is public or private relate to the extent to which we are ethically bound to seek informed consent from social media users.

There is also the issue of social media data containing data from people from broader networks, as in the case of people commenting on a social media user's post.

2. Informed consent

Informed consent is a critical component of the ethics of all types of research. In more traditional research approaches, informed consent is usually built in to the research design, for example in the form of consent forms or boxes to be ticked and signed on questionnaires. Social media-based research on the other hand presents problems concerning the informed consent of participants. In many cases, a social media user's data is accessed and analysed without informed consent having first been sought. 'Participants' in such research are rarely aware of their participation. Acquiring informed consent becomes more problematic the larger the data set, and can seem virtually impossible in aggregate data containing thousands or even hundreds of thousands of data units. Further, it is tempting to conflate a social media user having agreed to the terms of conditions of the platform (many of which include clauses on the accessing and re-use of data by third parties) with informed consent in research—problematic especially given that many social media users report not having read the terms and conditions properly. Important aspects of informed consent, such as the right to withdraw, are made more complicated in social media research (British Psychological Association, 2013) – for example, does deleting a post or account equate with a withdrawal from research, and is a researcher aware when this happens? When working with social media data, there are some conditions in which researchers will be more ethically bound to seek informed consent, such as when accessing data which social media users expect to be private (see above section).

3. Anonymity

Anonymity is a key consideration in research ethics, particularly in qualitative research practices or when data sets are shared outside of the original research team. Concerns over anonymity and online data are not new. With traditional forms of research, it is generally straightforward to anonymise data so that research participants cannot be identified. When working with social media data, however, anonymising data is more complex – anonymisation procedures are still evolving for aggregated or big data, and it is difficult to anonymise individual data extracts (such as Tweets) when these are reproduced in publications and during presentations.

This is further complicated when some platforms insist on units of data being republished only in their original form and attributed to the original poster. Different issues arise for different types of data too – the information contained within a text-based unit of data is different to what can be gleaned from images, audio- and video-format social media data. Given that social media companies tend to store data and metadata for long periods, and that much of this data is searchable, anonymisation in secondary use of data in some cases becomes challenging.

Further problems arise when data sets are exported to external coders and research partners. Issues of anonymisation become more critical in cases where data sets or individual units of data are published – for example online, in journal papers and at academic conferences.

Protecting the identity of unwitting participants becomes even more crucial when the data accessed refers to sensitive subject matter, particularly when exposing such data in new contexts and to new audiences may place the social media users at potential risk.

4. Risk of harm

Related to concerns over identity breaches is the risk of harm that researchers potentially place on their research subjects. The Association of Internet Researchers (2012) suggest **that a researcher's responsibility towards his or her participants increases with the increased risk of harm to those participants, or increased vulnerability of individuals or groups online.** This risk of harm is most likely where a social media user's privacy and anonymity have been breached, and is also greater when dealing with more sensitive data which when revealed to new audiences might expose a social media user to the risk of embarrassment, reputational damage, or prosecution (to name a few examples). This, of course, must be balanced with a **duty of care** on the part of the researcher to report concerns such as abusive or threatening behaviour online to the appropriate channels. It is not always clear to the researcher whether or not the data they have accessed, collected, analysed or reused can be retraced in its original online context, or what the repercussions of such retracing might be. Of particular concern is the republishing of quotes that have been taken from social media platforms and republished verbatim, as these can lead us, via search engines, straight back to their original location, often then exposing the identity and profile of the social media user they originate from (British Psychological Association, 2013). There may be issues in verifying information such as whether a participant is a child, or of sound-enough mind to understand the easily accessible nature of their data. This becomes of increased importance when dealing with sensitive or potentially embarrassing data. Therefore, where data deals with very sensitive topics, it becomes important to revisit the other concerns, ensuring that confidentiality and anonymity has been fully protected, and to consider whether or not to seek informed consent. Risk of harm might not be present in all instances in which a researcher wishes to cite social media data, for example when such data is shared by public bodies or organisations, or when the social media user is clearly aiming for broad readership (e.g. by using hashtags in Twitter).

5. Terms, conditions and legalities

Before **considering other aspects of the ethics of your social media project, it is vital to consult with all other relevant terms, conditions and guidelines.** Firstly, you should carefully read through all of the relevant terms and conditions of the platform(s) that you will be using to obtain your data. These terms and conditions will include those aimed at users, and might also include those aimed at third parties wishing to access data from the platform. Even if you have read the terms and conditions of a specific platform at some point, it is worth bearing in mind that the terms and conditions of social media platforms change regularly in accordance with changes made to the platform, or changes in how the platform owners wish to make profit from the platform.

Reading these documents may seem tedious – indeed, many platform users do not read them, and inadvertently agree to things they do not realise they have agreed to (such as use of their data by third parties!). But being familiar with the most current terms and conditions will protect you from potential legal action should you violate them.

Please note that jurisdictions within and outside the EU/EEA may have different rules for using the data. E.g. Twitter has different rules for US and non-US users. Other services have different conditions for different countries incorporated into one document.

Not only must the researcher be aware of the terms and conditions of using a particular platform, but must ensure that the participants are also fully aware of them and how they may impact on the protections of their data. It is therefore recommended that any use of particular platform must be fully disclosed in the **participation sheet and inform consent form**, for the participants to fully agree to all aspects of the research project, the use of a particular platform included.

You must ensure that the terms and conditions are not in conflict with the UOL research ethics policy or UOL data protection policy. Any query should be forwarded to the research services at research.ethics@sas.ac.uk.

6. Privacy and risk

You must determine whether the data you wish to access is really public, and if it is not to decide how - or indeed if - to proceed. The question as to whether to consider social media data as private or public comes down, to some extent, to whether or not the social media user can reasonably expect to be observed.

Things to consider here are:

- is the data you wish to access on an open forum or platform (such as on Twitter), or is it located within a closed or private group (e.g. within Facebook) or a closed discussion forum?
- Is the group or forum password protected?
- Would platform users expect other visitors to have similar interests or issues to themselves?
- Does the group have a gatekeeper (or admin) that you could turn to for approval and advice?
- How have users set up their security settings?
- Data accessed from open and public online locations such as Twitter present less ethical issues than data, which are found in closed or private online spaces. Similarly, data posted by public figures such as politicians, musicians and sportspeople on their public social media pages is less likely to be problematic because this data is intended to reach as wide an audience as possible.

If the data you wish to access is held within a group for which you would need to gain membership approval, or if the group is password protected, there are more ethical issues to take into consideration. Your first port of call should be to make contact with the site or group admin. They will have an understanding of the social dynamics of the group and will decide how to proceed. They may wish you to seek consent from individual group members for you to access their data, or offer group members the option to 'opt out' of the research (therefore you could use peoples' data unless they specify otherwise).

- Will you be asking questions of social media users in order to produce new data on a given subject?

If so, it is vital that you are transparent about your own identity (a researcher in a university) and that responses will be used as data in your research.

The 'blurring of boundaries' between researcher & participant is a further consideration - your own social media activity (or that of people you know) may be part of the dataset you are researching, which is potentially problematic. In this case care needs to be given to how such research methodologies and findings are reported. Also, the researcher themselves might become searchable by participants, **meaning that you should pay attention to your own online identity and privacy.**

- **Will you be dealing with young or vulnerable participants?** You must ensure that you have taken all possible precautions to rule out the use of data by vulnerable adults (i.e. those with additional educational needs) or children (or in the case of children, seeking parental consent).

Social media can often make it difficult to identify such individuals, not least because people often shield their true identities on social media platforms and discussion forums. Importantly, if data is suspected to originate from young or vulnerable individuals, informed consent cannot reliably be given so this data should be eliminated from the research.

- **Is the data potentially sensitive?** Is the data about fairly mundane daily activities or opinion, or is there the potential to cause harm to social media users should their data be exposed to new audiences?

Less sensitive data might include postings about, for example: the weather, recipes or consumer preferences. More sensitive data includes postings about, for example: criminal activity such as driving offences or the use of illegal drugs; financial problems; mental health issues and feelings of suicide; extramarital sexual activity; controversial political opinions and activism. **It is your responsibility to decide whether the content is sensitive and if so to determine an ethical way of working with the data, taking into account the UK Prevent Duty or the School's guidance on the matter of sensitive research.**

If there is risk of harm to individuals whose data you are using, you must either

a) paraphrase all data which is republished in research outputs, having taken steps to ensure that the paraphrased data does not lead interested parties to the individual's online profile;

b) seek informed consent from each person, should you wish to (or need to) use their data in its original form in research outputs

or c) consider using a more traditional research approach where consent and confidentiality can be more safely ensured.

It is also important to take these things into consideration in terms of whether you can share data sets (covered in more detail below).

There might be cases where it is not straightforward to seek consent. Conducting critical discourse analysis of harmful or ideological social media content (such as found in Neo-Nazi online groups) is one such. It may be useful to refer to our [guidance](https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/) which sets out what can be considered '**special category**' data: <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/> . **The researcher must also reflect as to whether it might be dangerous for them to get in touch with these social media users.** One may argue for such material as being exempt from the seeking of informed consent, in order to both protect the safety of the researcher, and to ensure that social media research ethics does not result in an indirect censorship of critical research. **However, the research project and methodology should be sent to the Research Ethics Committee for full review.**

7. Re-use and re-publication

There are different types of re-use or re-publication to be taken into consideration when you are working with social media data. You may wish to publish your research results in a number of different formats, for example:

- online blog posts;
- journal papers;
- conference presentations (including the submission of papers to conference proceedings);
- book chapters;
- articles online such as in The Conversation.

When reporting your findings,

- do you want to use units of data (such as individual tweets or Facebook postings, or Instagram images with corresponding text) to illustrate the themes that have arisen in the data?

If so, you need to review whether it is ethically sound to do so. You may feel that you need to either paraphrase the data or seek informed consent from individual platform users before you can do this.

- have you given consideration to the format of the data you are working with? – is it in photographic, audio or video formats? Are there copyright issues to be considered when re-publishing? Can such data compromise the anonymity of individuals or groups?

Sharing data sets is another consideration and something that is increasingly expected, for example by external funding bodies wishing to encourage a transparent and replicable research process. You should consider **whether it is ethically sound to share your data set, as well as checking the platform terms and conditions to determine whether they allow or prohibit it.** If it contains data that could cause harm if re-published, then either the sensitive data should be removed or paraphrased, or the data set should not be shared at all. In cases of aggregate data where the individual units (or postings) are no longer discernible, it is generally safe to share the data set. If the data set does not contain sensitive data, or if it is not possible to identify individuals based on the data set, it is also safe to share. If you are not happy about sharing the data, you will need to explain the reasons why when asked to upload or otherwise share the data set.

Special consideration needs to be given to anonymisation of social media users. **In almost all cases, it is important to ensure that users are anonymised in research outputs.** There are some exceptions to the rule, for example public figures and organisations seeking to share their data as widely as possible. Arguably, data that is not in any way sensitive (such as postings about the weather or consumer preferences) are unlikely to cause harm to individuals, therefore you can argue that it is unnecessary to anonymise content of this nature.

8. Are the data collected protected by IPR?

Individual records can be protected by copyright. The whole database or its parts can be protected by sui generis database rights. Some data are not protected at all and can be harvested at will. EU law allows certain statutory exceptions for non-commercial research purposes. Compiling, validating or structuring the data can give rise to a completely new IPR

owned by the consortium or its individual members. Adding content to the database usually requires consent from the IPR holder. If the new IPR was created in the phase of research, the decision to publish must not violate the rights of any co-author. Researchers should be aware of the citation requirements for the used in order to avoid a charge of academic misconduct. Note that there is variation in national laws.

Summary

Good ethical practices researchers can adopt for internet-based research are summarised below:

- Ensure transparency while recruitment, especially while recruiting using an acquired email list or through closed or open social media groups.
- Consider participants' expectations, perceptions, and awareness about privacy.
- Ensure compliance to national data protection laws and the applicable ethical codes.
- Use a password identity system to ensure authentic participants.
- Ensure that consent is informed and knowledgeable.
- Ensure that a privacy statement is provided which explicitly mentions purposes for which the data will be used, participants' rights, and strategies adopted by the research team for maintaining anonymity, confidentiality, and data protection.
- Assign each participant a unique identification number/code to track progress and maintain security.
- Provide opportunities to participants to reach the research team or clarify any concerns.
- Honour participants' request to delete their records.
- Ensure encryption and Secured Socket Layer protection systems.
- Use password protected computer directory to store the data.
- Remove all identifiers before data analysis.
- Ensure no unauthorized person is able to access the data.
- Ensure that data are not used for subsequent non-research purposes.

Remember, everything done online will have a digital print.

Further reading and Acknowledgement

This document was written based on guidance that is available online. They are listed below:

[Facebook as a Research Tool for the Social Sciences](#), published in the September 2015 issue of *American Psychologist*. , <http://psycnet.apa.org/journals/amp/70/6/543>

Facebook for researchers: <https://phdlife.warwick.ac.uk/2014/01/03/facebook-for-researchers/>

Social Media, the GDPR and Data: 18 Key Things You Need to Know : <https://dma.org.uk/uploads/misc/social-media,-the-gdpr-and-data1.pdf>

GDPR (<https://ico.org.uk/media/for-organisations/guide-to-the-general-data-protection-regulation-gdpr-1-0.pdf>)

Skype: An Appropriate Method of Data Collection for Qualitative Interviews?: <https://pdfs.semanticscholar.org/a38b/5c503a48517a8a2644e561ac59e8f16e3afa.pdf>

Twitter as a Data Source: An Overview of Ethical, Legal, and Methodological Challenges.: http://eprints.whiterose.ac.uk/126729/8/Normal_-_Ethics_Book_Chapter_WA_PB_GD_Peer_Review_comments_implemented_1_.pdf

'*Online interviewing with interpreters in humanitarian contexts*' is a research paper written by Anna Chimento, Laura Machin, Atif Rahman and Lucy Frith in the International Journal of Qualitative Studies on Health and Well-being. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5935183/>

Further reading and resources Association of Internet Research (2012). Ethical Decision-Making and Internet Research. Available at: <http://aoir.org/reports/ethics2.pdf>

Social Media Research: A Guide to Ethics :
https://www.gla.ac.uk/media/Media_487729_smxx.pdf

British Psychological Association (2013). Ethics Guidelines for Internet-Mediated Research. Report available at: <http://www.bps.org.uk/system/files/Public%20files/inf206-guidelinesfor-internet-mediated-research.pdf>

CASRO (2011) Social Media Research Guidelines. Available at:
http://c.ymcdn.com/sites/www.casro.org/resource/resmgr/docs/social_media_research_guidel.pdf

Collaborative Online Social Media Observatory (COSMOS) Ethics Resource Guide. Available at: <https://www.cs.cf.ac.uk/cosmos/ethics-resource-guide/>

ESOMAR (2011) Esomar Guideline on Social Media Research. Available at:
<https://www.esomar.org/uploads/public/knowledge-and-standards/codes-andguidelines/ESOMAR-Guideline-on-Social-Media-Research.pdf>

Evans, H., Ginnis, S. & Bartlett, J. (2015) #SocialEthics: a guide to embedding ethics in social media research. Report available at:
<https://www.ipsosmori.com/Assets/Docs/Publications/im-demos-social-ethics-in-social-media-researchsummary.pdf>

Fossheim, H. & Ingierd, H. (2015). Internet Research Ethics. Available at:
<https://press.nordicopenaccess.no/index.php/noasp/catalog/view/3/1/9-1>

Natcen (2014) Research using Social Media: Users' Views. Available at:
<http://www.natcen.ac.uk/media/282288/p0639-research-using-social-media-report-final190214.pdf>

Salmons, J. (2014) New Social Media, New Social Science... And New Ethical Issues! Report available at: <https://drive.google.com/file/d/0B1-gmLw9jo6fLTQ5X0oyeE1aRiQ/edit>

7th August 2020.