INTERNET-MEDIATED RESEARCH

Contents

1. Introduction

Internet-mediated research can be broadly defined as *"any research involving the remote acquisition of data from or about human participants using the internet and its associated technologies"* (British Psychological Society, 2017)[1]. Technological developments (the use of social media, access to the internet through phones and other devices, big data) have led to research opportunities but also complexities when it comes to assessing and addressing the ethical issues. A discussion of internet research is provided within the UK Research Integrity Office's guidance on Internet-mediated Research (2016)[2].

While the principles of maximising benefit and reducing harm; respecting the rights and dignity of individuals; obtaining informed consent; and conducting research with integrity apply to internet-mediated research as they do to other types of research, how these are applied will depend on the

[1] The British Psychological Society's Ethics Guidelines for Internet-mediated Research (2017)
[2] http://ukrio.org/wp-content/uploads/UKRIO-Guidance-Note-Internet-Mediated-Research-v1.0.pdf

context of the research and many decisions will need to be made on a case by case basis. Measures taken should be proportional to the level of risk and to the potential harm to participants. Conflicts of interest between the rights of individuals, the overall benefits of the research and the terms of use of the platform will need to be identified and addressed.

Rather than viewing the consideration of the ethical issues as a one-off process, consideration should be given to the ethical issues associated with internet research during all stages of the research project, from the design stages to data collection, analysis and dissemination of the results. The Association of Internet Research (AoIR)[3] has produced a set of questions for researchers and the BPS Ethics Guidelines for Internet-mediated Research[4] contains a table summarising the main ethical issues for researchers to consider.

This guidance should be read in conjunction with discipline-specific principles and frameworks in order to highlight challenges and to inform decision-making about ethical issues associated with internet research. It is helpful to bear in mind that there may be more than one approach to addressing the ethical issues; ethical pluralism and cross-cultural awareness are important, particularly for internet research involving participants or data from a range of countries and cultures.

This document discusses how some of the fundamental principles of research ethics apply to internet-mediated research including: respecting those involved, expectations around privacy, data management, obtaining informed consent and protecting the researcher(s). Additional guidance is provided for specific contexts**: online surveys, mobile phones and apps, using quotations in observational studies, terms and conditions of internet platforms, mass-emailing** and **research involving the Dark Web**.

Internet-mediated research can be additionally complicated by the ways consent can be interpreted: "A Guide to Unobtrusive Methods in Digital Ethnography"[5] provides some conceptual guidelines for "lurking", taking part in digital communities, public vs. private data, etc. See for example Section 6 (Publishing tweets and other quotations from social media).

2. *Respecting individuals and communities*

Consideration should be given to whether the research involves any ethically significant risks for the individuals or (online) communities involved and take proportionate steps to address these.

Particular care may be needed if the participants could be considered as vulnerable or the research could involve sensitive topics, i.e., participants' sexual behaviour, their illegal or political behaviour, their experience of violence, abuse or exploitation, their mental health, or their gender or ethnic status (from the ESRC explanation of sensitive); in such cases researchers may find it helpful to seek advice either from within the Department or from their ethics committee.

Researchers should consider the extent to which the online and offline identities of the participants may differ. The steps to be taken if the age of the participants is unclear should be considered on a case-by-case basis, bearing in mind the sensitivity of the topic as well as the design of the research.

[3] https://aoir.org/reports/ethics3.pdf
[4] The British Psychological Society's Ethics Guidelines for Internet-mediated Research (2017)
[5] Ugoretz, Kaitlyn. (2017). A Guide to Unobtrusive Methods in Digital Ethnography

The more vulnerable the participant the greater the obligation to protect them (cf. AOIR 1.0, p.5; Tiidenberg, 2018[6]). If the data is likely to include data relating to individuals who are vulnerable, e.g. because of their age or because of the nature of the research project, extra care may be needed. Please also refer to the University Guidance and Code of Practice on Safeguarding.

As with other types of research, it is expected that consent from a parent or legal guardian will be obtained in internet-mediated research which recruits young people, though exceptions may be made for youths classing as 'Competent Youths' (see related guidance on this topic in the Frequently Asked Questions (C12) section of the CUREC website as well as the Best Practice Guidance on Research with Competent Youths. For some research, offline processes for obtaining parental/ guardian consent before conducting research with children or adults at risk may be appropriate.

For advice on writing in an accessible way for participants, which is particularly important if there is not going to be any in-person interaction with them, refer to the guidance available via the Research Support website.

It may be helpful to distinguish between types of people whose data will be used, e.g., people using the Internet for everyday purposes (such as communicating or socialising) who may not be expecting to be part of a research study, research participants who have been specifically recruited to take part in a particular project, public figures whose online presence/ postings are intended to be public and microworkers, e.g. "Mechanical Turks", when assessing ethical issues such as informed consent and privacy.

See also Section 12, Research involving online communities (e-sports/gaming, special interests, etc.)

3. Public/ private distinction

The status of different sources of data obtained via the internet may not always be clear-cut and this has implications both for determining an appropriate level of consent and for deciding how much to protect individuals' identities.

"*Not all information openly available online is public, and thereby [can] be made an object of research without informing and obtaining consent from those concerned. Nor can all information that is of a public nature be indiscriminately used for research purposes.*"[7]

The likely expectations of individuals should be respected. Individuals may consider their posts to be private even if they are posted in a public forum. It may be possible to combine and analyse data sets to reveal characteristics or even identities.

On the other hand, researchers should bear in mind that in some cases not disclosing the identity of individuals could be misinterpreted or cause offence.

The type and accessibility of the platform, the form of communication, whether there are access restrictions, e.g., age limits, the number of users or cultural or social requirements, may be helpful

[6] Tiidenberg, Katrin. (2018). Ethics in Digital Research

[7] The Norwegian National Research Ethics Committees' Guide to Internet Research Ethics (2019)

factors to bear in mind when determining whether the data should be considered public, and the steps needed to obtain consent[8]. A forum may have been created as a safe space for specific topics.

Extra care may need to be taken when managing potentially sensitive data, particularly if informed consent has not been obtained. Researchers should be sensitive to cultural, individual, and role differences in researching groups of people with distinctive characteristics. Laws and cultural norms in relation to privacy should also be respected.

Even if the data are publicly available, several ethical concerns remain, e.g.:

- the post/ data must not be misrepresented by the researcher;
- the user's data must not be 'triangulated' in such a way that the researcher reveals identifying or potentially harmful information that the user did not originally intend to share (e.g. linking Twitter posts with information on another platform by the same user);
- care should be taken with inferring personal information from the information provided (e.g. gender from screen names, geographical location using analysis of text or photos);
- the poster may not be who they say they are (for example, they may be a minor, or using someone else's account), or in a state of reduced capacity to consent when they make the post;
- changes to the data set over time, e.g., changes to the privacy settings on the site, removal of data (and whether and how data removed from a site should be removed from a previously collected dataset);
- limits to the anonymisation of data, especially qualitative data;
- the ease with which data could be traced back to individuals or online communities, particularly from research outputs. E.g., when using a direct/ paraphrased quotation.

4. Confidentiality and security of online data

Guidance on research data management is available within CUREC's Best Practice Guidance 09 on Data Collection, Protection and Management and within the Research Support guidance on working remotely with participant data. Researchers should follow the data protection standards and legislation of the countries where the research is being undertaken. If the location of a participant is unknown, then the UK and the University of Oxford's requirements must be followed.

The risks associated with identification should be assessed within the context of the research. Researchers should clarify whether there could be additional risks of harm to the participants from their data being used for the research project or quoted in any research outputs. It may be significantly easier to identify individuals from internet-mediated research than from other types of research, for example by searching online for a phrase from a quotation. Note that UK and EU data protection laws specify that e.g., usernames are considered personal data.

"While much internet communication is often effectively public through greater visibility, traceability and permanence, it is not always apparent whether this makes it ethically acceptable to use such data freely for research purposes." (BPS 2017[9], page 7)

Privacy and confidentiality of data may also be more difficult to manage in internet-mediated research because researchers are not in control of online communication networks, increasing the

[8] https://aoir.org/reports/ethics3.pdf
[9] https://www.bps.org.uk/news-and-policy/ethics-guidelines-internet-mediated-research-2017

risk of third-party interceptions. As in non-online studies, researchers must not make misleading or inaccurate statements about privacy or confidentiality in consent processes.

As part of information-giving prior to seeking consent, it is good practice to clarify any limits to confidentiality or anonymity, e.g.:

- General: '*Although every reasonable effort has been taken, confidentiality during actual internet communication procedures cannot be guaranteed*'.
- For research using third party websites to administer surveys: '*Data may be stored on backups or server logs beyond the timeframe of this research project'*.
- When seeking participants' informed consent for interviews conducted in writing (email, text message, chat forum, etc.) participants should be made aware of limitations to the security of the communication, e.g., risks of other people gaining access through a data breach, hacking or lawful access. : '*Whilst the researcher will take every care with the data you share, you should only take part in the study if you/ your company are prepared for your responses to be made public. This is because the record of the conversation could theoretically be breached or hacked, outside the control of the researcher*.'

Both the use of metadata and cross-referencing carry a greater risk of privacy breaches for individuals and could affect their autonomy over their online information.

In less clear-cut situations, it may be worthwhile to take into consideration the views of moderators or other gatekeepers of the platform in respect to how identifiable individuals or the platform should be.

Consider also whether other individuals could be identifiable from the data, e.g., friends, followers or others mentioned by or connected to the participants, and how to address the associated ethical issues.

Some platforms, e.g., Twitter, have conditions relating to the use of direct quotations in their terms of use. It is the researcher's responsibility to check the relevant social media platform's terms and conditions. Be aware that terms and conditions may be based on the laws of a country outside of the UK, and that these may change during the research. If researchers do expect to breach any terms and conditions, this should be explicitly addressed in the research ethics application (see section a) below).

a) Legal implications

Consideration should be given to whether the data set could change, particularly to comply with the GDPR and UK Data Protection Act's right to be forgotten[10].

Consideration should also be given to whether it may be necessary for the researcher to report harm, risk of harm, or criminal activities (if this is something that could arise during the research) and, if so, what the limits to the researcher's obligations might be.

Generally, researchers should check, when planning their research, that they will not contravene the terms and conditions of the platforms or apps they are studying or using to conduct their research.

[10] https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-erasure

Ethics applications for research where the researcher(s) intends to breach the terms of an internet platform will be assessed on a case-by-case basis and should address the following:

i. whether there are different ways of conducting this research to avoid the breach of contract;
ii. whether similar types of research have been conducted previously (and with what justification);
iii. the degree of public interest in the research;
iv. the likely public benefit of the research;
v. the level of experience and supervision of the researcher(s);
vi. the degree of potential harm to participants in the research (including any possible re-identification of participants in the research);
vii. the degree of potential harm to the researcher(s) in conducting the research;
viii. the degree of potential risk associated with the processing of personal data associated with the research (and any possible breaches of data protection legislation);
ix. the potential reputational risk to the Department and the University from the breach of contract. Bear in mind that it may be necessary to seek advice from Legal Services and the Chair of CUREC (and potentially the Registrar) as appropriate, depending on the level of risk;
x. how the results of any such research should be published or publicised (and especially whether verbatim quotation or other approaches in the publication might increase risk to the platform users).

5. Obtaining informed consent

A proportionate approach should be taken when deciding an appropriate form of informed consent for the research. In addition to the guidance on obtaining informed consent and communicating with participants, it may be helpful to consider the following:

- The sensitivity of the research (the ESRC gives the following examples of sensitive topics: participants' sexual behaviour, their illegal or political behaviour, their experience of violence, abuse or exploitation, their mental health, or their gender or ethnic status);
- Reasonable expectations relating to privacy for the research setting;
- Whether it could be possible to identify the individuals, both from the data and from any research outputs.
- It is possible that your sample may include minors or people with reduced capacity to give consent, and this might not be apparent; you may need to consider appropriate and proportionate measures to handle this.

If researchers are not planning to obtain participants' informed consent this must be explained and justified within the ethics application.

A proportionate approach should also be given to allowing participants the ability to withdraw from the research, both during and after participation. Where participants interact with online research materials or researchers themselves to generate fresh research data, participants should be made aware of how to withdraw themselves and/ or their data both during and after data collection. Participants should be informed of any time limits as to when they can withdraw data they have provided (e.g., after data has been anonymised).

Consider whether it is necessary to share or to publish identifiable data. If identifiable data is to be shared with others or published it is good practice to obtain participants' informed consent. If it is not possible to obtain informed consent, e.g., if contacting the individuals cannot be done or the dataset is very large, it is good practice to de-identify/ pseudonymise the data. Checks need to be

made that the strategies used to pseudonymise data are sufficiently effective. For example, simply changing a username does not prevent a comment being found via search engines.

Care needs to be taken if researchers have existing relationships with participants. Befriending, following or connecting with research participants is not considered good practice and may go against the terms of use of the platform. If there are good reasons to connect with participants via social media, it is advisable to do so via an account that has been set up specifically for the research rather than from a researcher's personal social media account. Zimmer (2010)[11] contains a discussion of this.

Researchers should always check whether material they wish to use is protected by copyright law or if there are conditions associated with its use.

Any deception of participants or withholding of information (including how the research will be used or published) should be explained and justified within the ethics application. Sufficient information must be provided within the ethics application to enable the ethics committee to make an effective ethics decision on whether it is reasonable for the information to be withheld and whether the consent process used is appropriate. If the information provided or disclosed to participants is limited or restricted to the extent that they are unable to make an informed decision about participating, an application should be submitted using the appropriate form for review by the IDREC rather than a CUREC 1, 1A or minimal risk application. Further guidance is available within CUREC's Approved Procedure 07 Deception of Adult Participants and within the BPS guidance[12].

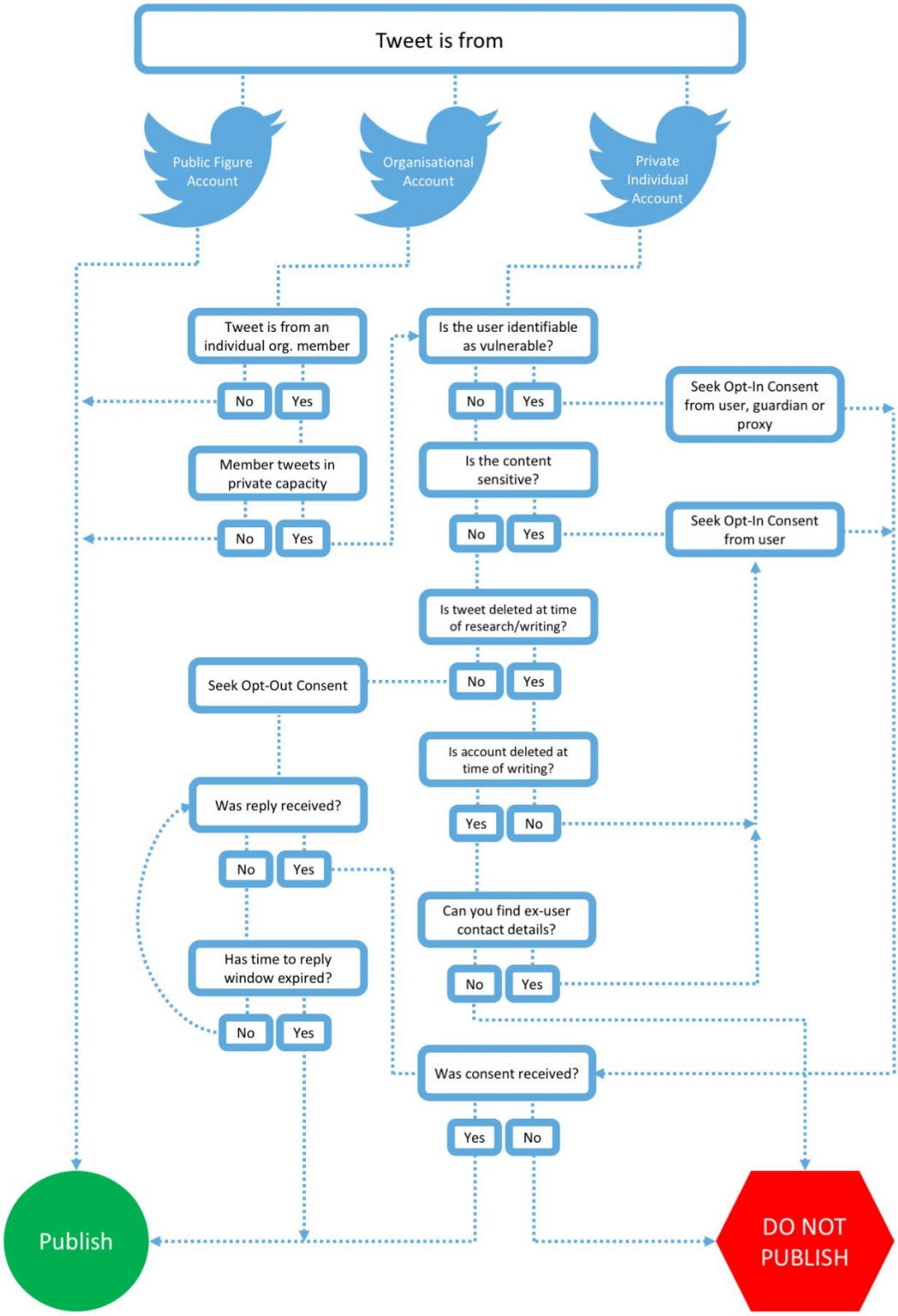
6. Publishing tweets or other quotations from social media sites

The following flowchart about reporting tweets and opt-out processes developed by Williams, Burnap and Sloan (2017[13]) may be a helpful guide (text version below):

[11] Zimmer, M. (2010) '"But the data is already public": on the ethics of research in Facebook', *Ethics and Information Technology*, 12 (4): 313–25 (accessed 5 November 2019

[12] The British Psychological Society's Ethics Guidelines for Internet-mediated Research (2017)

[13] Williams M.L., Burnap, P. and Sloan, L. (2017) 'Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users' Views', Online Context and Algorithmic Estimation. *Sociology,* 51 (6) 1-20.


Tweet is from
Public Figure Account
Organisational Account
Private Individual Account
Tweet is from an individual org. member
No
Yes
Member tweets in private capacity
No
Yes
Is the user identifiable as vulnerable?
No
Yes
Seek Opt-In Consent from user, guardian or proxy
Is the content sensitive?
No
Yes
Seek Opt-In Consent from user
Is tweet deleted at time of research/writing?
Seek Opt-Out Consent
No
Yes
Is account deleted at time of writing?
Was reply received?
Yes
No
No
Yes
Can you find ex-user contact details?
Has time to reply window expired?
No
Yes
No
Yes
Was consent received?
Yes
No
Publish
DO NOT PUBLISH

Text version of flowchart:

Researchers may **publish** a tweet if one or more of the following applies:

- The tweet comes from a public figure account;
- The tweet comes from an organisational account – either an organisational account that isn't that of an individual organisation member, or the account of an individual organisation member who is not tweeting in a private capacity;
- Opt-out has been deemed necessary, and consent has been given, or no reply has been received and the window for replying has expired;
- Opt-in consent has been deemed necessary, and consent has been given.

Researchers should **seek opt-out permission to publish** a tweet from a private account, or an organisation account where an individual organisation member is tweeting in a private capacity, as long as all of the following apply:

- The user is not identifiable as vulnerable;
- The content of the tweet is not sensitive;
- The tweet has not been deleted at the time of research/ writing.

Researchers should **seek opt-in consent to publish** a tweet from a private account, or an organisation account where an individual organisation member is tweeting in a private capacity, if one or more of the following applies:

- The user is identifiable as vulnerable;
- The content of the tweet is sensitive;
- The tweet has been deleted but the account has not been deleted;
- The tweet and account have been deleted, but the researcher can find the ex-user's contact details.

Researchers **should not publish** if one or more of the following applies:

- Opt-in consent has been sought, and not received;
- Opt-out permission has been sought, and permission has been refused;
- The tweet and account have been deleted, and the researcher cannot find the ex-user's contact details.

These principles should be followed for other social media platforms, with similar terms of service allowing re-use of posts for academic research.

7. Protecting the researcher

Consider whether the research could involve security-sensitive material or lead to online or offline threats or harassment of the researcher(s), e.g., stalking, harassment, distributed denial-of-service (DDoS) attacks, identity theft, etc. The University Information Security team is happy to be contacted for advice: if needed, they can conduct a security audit of your internet footprint and/ or monitor your University accounts.

Researchers can also be at psychological risk from exposure to unpleasant content or individuals, or a want to talk about their experience but feel bound by confidentiality requirements. The University Counselling Service is always available; Departmental RECs may also be able to direct you to more informal support networks. The Social Sciences Division has a factsheet and holds workshops to help address emotional impact and vicarious trauma.

There may also be legal risks. Guidance for researchers whose research may potentially fall within the scope of the 'Prevent' duty is available within CUREC's Best Practice Guidance 07 on the Prevent Duty.

8. *Online surveys – additional considerations*

Data protection requirements with respect to UK GDPR and the Data Protection Act must be followed. In accordance with the principle of data minimisation, researchers should only collect personal data that they need for their research. For example, participants could be asked to indicate their age by selecting from age groups, or by providing their year of birth rather than providing their exact date of birth. In some cases, it may be possible to conduct a survey without collecting any personal data at all.

Researchers are encouraged to use a platform where the data is stored within the UK or EEA. Jisc Online Surveys is the University's recommended platform for online surveys. IT Services manage an organisation-wide licence for Jisc Online Surveys and any student, staff member or academic visitor can request an account for creating surveys. It is not possible to access information about respondents' IP addresses if Jisc Online Surveys has been used.

As well as Jisc Online Surveys, the Information Security team has approved the use of the online survey platforms Microsoft Forms (the version within Nexus365), Qualtrics and SmartSurvey for collection of all types of personal data, including confidential data. SurveyMonkey and Google Forms are not recommended for collecting personal data.

Researchers should also ensure that online surveys are set not to collect IP addresses unless these are needed for the research. Certain online survey platforms (e.g., Qualtrics) include the IP addresses of respondents as part of the survey results by default, but this can usually be switched off for individual surveys.

Researchers should take advice from Information Security about other third party providers who may collect or process personal data on their behalf as a third party security assessment (TPSA) may need to be conducted.

Guidance on payments to participants is available within CUREC's Best Practice Guidance (05) on Payments and Incentives in Research.

a) *Implied Consent and Informed Consent in Online Surveys*

For research with more straightforward ethical issues, such as completion of a simple online questionnaire, completion and submission of the questionnaire can imply that consent for the use of the questionnaire data has been granted. However, the questionnaire should be preceded by written information about the research and its aims (including information about how the data will be stored and published and a tick box confirming that participants meet the inclusion criteria and agree to take part). Further guidance and a suggested format for the participant information is available via the Research Support website. This template should be used as the basis for the information provided to respondents in all online research surveys.

The participant information should explain how participants can withdraw, e.g., by closing the tab in their browser or by clicking on a 'withdraw' button within the survey itself that leads to a debrief page. Whether or not the data (including IP addresses) will be retained must be made clear. Any limits to withdrawing, for example once the data has been anonymised, should be explained to the

participants. It is normally good practice to give participants the option of not answering questions within a survey, either by not requiring an answer or by including a "prefer not to say" option. Participants should be able to contact the researcher(s) if they have questions, concerns or would like to provide feedback on the survey.

The Information Compliance website has guidance on creating privacy statements and some templates. This information must be included within the participant information or a link to a separate privacy notice must be provided.

b) The Data Controller and Data Processor for online surveys

To determine the Data Controller and the Data Processor in a particular scenario, a helpful starting point is to establish who is making key decisions on what personal data to collect and how this data is processed. It is the answer to this question that determines who the data controller is, and who is the processor in a given situation. In most cases the Data Controller will be "The University of Oxford". It is possible to have more than one Data Controller, e.g., if there are collaborators at another institution. Colleges are separate legal entities so have different Data Controllers. Further guidance is available at https://researchsupport.admin.ox.ac.uk/policy/data/responsibilities.

If a University of Oxford researcher were to collect personal data on behalf of another organisation and the other organisation was providing full instructions as to how they want the personal data to be processed, it is likely that the University of Oxford would be seen as the Data Processor. This can happen in a research project which is sponsored by another university, when the other institution is leading the research and requests Oxford researchers perform a particular task to assist. If the University of Oxford simply follows instructions on what personal data to collect and process, it would be acting as a Data Processor. Importantly, this would be the case even if the other organisation never had direct access to the personal data that Oxford researchers have collected and processed on their behalf. Similarly, if Oxford researchers instruct an online survey provider to collect or process personal data on their behalf, it is likely that the University of Oxford will be the data controller even if the Oxford researchers are only given personal data in an anonymised form.

Note that anonymisation is very difficult to achieve to UK GDPR standards and goes a long way beyond the straightforward removal of personal identifiers. Further guidance on the challenges of anonymisation is available here: https://researchsupport.admin.ox.ac.uk/policy/data/scope.

9. Conducting interviews online

Microsoft Teams is the University's approved platform for virtual meetings and currently the only platform approved for conducting meetings where confidential or sensitive subjects will be discussed. Practical considerations when planning interviews to be conducted online include taking into account varying degrees of digital literacy and access to technology. It may be more difficult to tell if the location is suitable, i.e. whether the participant is in a safe place or if they could be overheard. Ideally this should be directly addressed through the process of obtaining participants' informed consent. Researchers must also consider the arrangement at their end. If the interviewer is working at home, for example, are they somewhere private where they are unlikely to be interrupted, with a suitable background? The physical separation between the interviewer and interviewee may also affect communication. It might be harder to tell if a participant needs a break or is upset when conducting interviews remotely. Checking if a participant is alright to continue part-way through the interview might be appropriate, particularly where sensitive issues are being discussed. Further guidance is available within CUREC's Best Practice Guidance (10) on Conducting Research Interviews and within the guidance for researchers working remotely with participant data.

10. Mobile phones and apps – additional considerations

Sensitive data from mobile internet connections (e.g., users' location and contact details stored on smart phones and tablets, as well as the metadata of their communications) raise additional ethical issues – especially as there may be a possibility of re-identifying cases in 'anonymised' datasets. For a broad discussion of these issues and helpful practical and ethical guidelines for researchers using datasets constructed with information from mobile devices refer to Ethical Privacy Guidelines for Mobile Connectivity Measurements (2013)[14]. Specific ethical and technical recommendations can be found in Best Practice Guidance (12) on Mobile App Design.

11. Emailing a large number of people

The University's IT Regulations state that "*Users are not permitted to use university IT or network facilities for… transmission, without proper authorisation, of e-mail to a large number of recipients, unless those recipients have indicated an interest in receiving such e-mail.*" "Large" refers to the number of individuals concerned, the volume of data, the variety of data, the duration of processing and the geographical extent of processing. Further guidance is available from the Information Compliance webpages.

Given the potential also for breaches of data protection legislation when constructing email mailing lists, researchers are advised to seek further advice from the University's Information Compliance Team and Information Security Team.

12. Research involving online communities (e-sports/ gaming, special interests, etc.)

Considerations in this area will depend heavily on the venue and type of research being conducted, but could include:

- Responsible engagement with gaming platforms, to get permission to conduct research.
- Consent: the "lurker" problem, or related issues where a research may be playing a game/ participating in the community while also conducting the research (see for example: "A Guide to Unobtrusive Methods in Digital Ethnography"[15]).
- The fact that children often play or participate in these online communities and might not disclose their age accurately: adequate and proportionate safeguards should be taken. Researchers are encouraged to seek advice from their DREC or IDREC if this is likely.
- Potential risk from toxic culture (verbal, written, or visual "banter", aggression, abuse, etc.) – this could create direct or vicarious trauma for the researcher (see Section 7, Protecting the researcher, above).
- Potential risk of direct online or offline abuse like stalking, doxing and SWATing (see Section 7, Protecting the researcher, above).
- Navigating the culture of the community you are researching: for example, understanding the local etiquette for types of conversation suitable "in game" vs. in a side-discussion in Discord; gaining membership to invite-only communities; following the meaning of in-jokes and memes; understanding when it is or is not appropriate to link people's online personas to their offline identity (or identities on other platforms), etc. Ethnographic research best practice is generally a good guide but it can be easy to underestimate the differences in sub-

[14] Zevenbergen, Bendert and Brown, Ian and Wright, Joss and Erdos, David, Ethical Privacy Guidelines for Mobile Connectivity Measurements (November 7, 2013)
[15] Ugoretz, Kaitlyn. (2017). A Guide to Unobtrusive Methods in Digital Ethnography

culture that may apply online, or the ways in which online behaviour can diverge from behaviour in real life. See Section 2, Respecting individuals and communities.

13. Research involving the dark web

The dark web forms a small part of the deep web, the area of the World Wide Web that is not indexed by web search engines. The dark web exists within overlay networks (darknets) that use the Internet but require special network-routing software, configurations, or permissions to access (most commonly, the TOR browser). Through the dark web, communications can be made by private networks and business conducted anonymously without divulging identifying information, such as a user's location. The dark web works much like the open web, and is used for many legitimate purposes, but it is also a home for criminal and other unsavoury activities due to the increased privacy.

If researchers plan to use the dark web as part of their research study, they are advised to seek advice from the relevant CUREC sub-committee or their DREC: potential risk will depend heavily on the type and design of the research activity. Researchers could encounter illegal, radicalising or deeply disturbing material, or come across personal information (especially in criminal data dumps) which falls under data privacy legislation. Researchers could be at risk of online or offline retaliation from criminals if they publish a paper exposing criminal activity, or at risk of travel bans or government legal action if they are reporting on banned political speech.

Researchers should consider the following before accessing the dark web:

1. the University **IT regulations**, which stipulate (inter alia) that

   "7. Users are not permitted to use university IT or network facilities for any of the following:

   i. any unlawful activity;
   ii. the creation, transmission, storage, downloading, or display of any offensive, obscene, indecent, or menacing images, data, or other material, or any data capable of being resolved into such images or material, except in the case of the use of the facilities for properly supervised research purposes when that use is lawful and when the user has obtained prior written authority for the particular activity from the head of his or her department or the chair of his or her faculty board (or, if the user is the head of a department or the chair of a faculty board, from the head of his or her division);
   iii. with the intention of drawing people into terrorism (contrary to the University's statutory duty under Prevent). […]"

   (These regulations are a reminder not to engage in unlawful activity online. However, it should be noted that the University cannot protect its staff or students from police/ security services action.)

2. the **researcher's supervisor** (and possibly **Head of Department** or other devolved authority)
3. the **University's Information Security Policy and Guidance**, for advice on how to put in place appropriate security measures if accessing sensitive material (or material blocked on the University network)[16]

[16] University of Oxford's Information Security guidance and policy

4. **Emotional Impact/ Vicarious Trauma** information and guidance (particularly if accessing material could potentially cause distress to the researcher(s)[17]
5. **CUREC 'Prevent' guidelines** , if there is a risk that the research topic could potentially come within the scope of the 'Prevent' duty, which seeks to prevent people from being drawn into terrorism[18].

a) Deepfakes

Deepfake video, audio and still images (where Artificial Intelligence (AI) is used to manipulate multimedia content to present an event or artefact that that did not actually occur) should be treated the same as other types of personal data, meaning that any analysis of them (beyond metadata) need to follow comparable rules (e.g., consent, paraphrasing, non-disclosure) to other visual/ audio data used in internet-mediated research. Appropriate care should be taken given that deepfakesmay have been made without consent or with the intention to harm others/ tarnish reputations. Researchers in some fields may need to be alert to the presence of unlabelled deepfake content masquerading as reality.

If your research may involve *creating* deepfakes (of humans or anything else; whether through publicly available generators, or developing your own generator), please consult your local Ethics Committee. The Computer Science DREC has some experience with these and can provide access to self-assessment tools or subject expertise if needed.

14. Further reading

All last accessed January 2021:

- Association of Internet Research Ethics guidance.
- Jane Bainbridge (2015) 'Call for better ethical standards in social media research' and Academy of Social Sciences, Summary of Conference on 'Ethical issues in social science research on social media', March 2016.
- Eynon, R., Fry, J. and Schroeder, R., 2017. The ethics of online research. The SAGE handbook of online research methods, pp.19-37.
- The British Psychological Society's Ethics Guidelines for Internet-mediated Research (2017).
- The Economic and Social Research Council's guidance.
- The Economic and Social Research Council's guidance on research with potentially vulnerable people.
- The International Visual Sociology Association's Code of research ethics.
- The Norwegian National Research Ethics Committees' Guide to Internet Research Ethics (2019).
- The University's IT Regulations.
- Tiidenberg, K. (2018). Ethics in digital research.
- UK Research Integrity Office's Good practice in research: Internet-mediated research (2016)
- Williams M.L., Burnap, P. and Sloan, L. (2017) 'Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users' Views', Online Context and Algorithmic Estimation. Sociology, 51 (6) 1-20.

[17] University of Oxford's Emotional Impact/ Vicarious Trauma information

[18] See CUREC's Best Practice Guidance on Prevent Duty

- See Zimmer, M. (2010) '"But the data is already public": on the ethics of research in Facebook', Ethics and Information Technology, 12 (4): 313–25.

15. Change history

Version No.	Description of changes	Previous version No.
7.1	New section on research involving online communities and revisions to the existing section on research involving the dark web.	7.0
7.0	With the exception of the section on the Dark Web and a few paragraphs and sentences elsewhere, the entire document has been rewritten and restructured.	6.5
6.5	Removed Appendix A – Template Information and Consent – to make separate template document. Removed two tables from section 3.1	6.4
6.4	Document made accessible for use as a pdf	6.3
6.3	Guidance updated following Information Security's approval of Microsoft Forms for online surveys	6.2
6.2	Guidance updated following Information Security's approval of Qualtrics for online surveys; Tweet flowchart added to the Informed Consent in Social Media Research section.	6.1
6.1	Updated to reflect the UK departure from the EU	6.0
6.0	Major changes to the sections about 'Deepfakes', 'Higher Risk Research' and 'Legal and Compliance Issues'; Clarifications on issues around public accessibility settings in social media platforms, and issues around analysing rare events that may endanger social media users' anonymity; Addition of new sections 'Terms and conditions of social media platforms or social networking apps' and 'Copyright'; Removal of the section about Safe Harbour/ Privacy Shield (essential brief information having been moved to footnote 1); Text from the previous 'recruitment' section split into more relevant sections; Additional Guidance text on the preferred online survey suppliers Jisc and RedCap; Revisions to Appendix A – Template Information and Consent; General text update.	5.3