

Guidance on the ethics of research using human participant data from on-line sources/social media

Please note that this guidance will be subject to regular review and updates by the Research Ethics Integrity and Governance team: always access University ethics guidance via the [University Ethics A-Z web-page](#) rather than as a saved downloaded document.

1. OVERVIEW AND PURPOSE

The purpose of this document is to provide guidance for University of Sussex staff and students planning to collect human participant data from on-line social media platforms (referred to here on in as social media research data.) This guidance does not provide detail on specific platforms – which are varied and constantly changing – but instead outlines the ethics and compliance issues that researchers should consider when designing a research project and provides advice on how to approach the ethics application and review.

The guidance is in three sections:

- Key ethical principles
- Data protection considerations and legislation
- Mass data

The following key documents have been consulted in the development of this guidance, and researchers are advised to refer to these when considering a research project involving social media data (please see the *References* list for full details):

- [University of Sussex Research Ethics Policy](#)
- [University of Sussex Research Ethics Policy Supporting Notes](#)
- [BPS Ethics guidelines for internet-mediated research](#)

This guidance should be used alongside existing University policy and guidance on ethics, research data management, and data protection.

[Policies at the University : Information Management : Governance and Compliance : University of Sussex](#)

2. SCOPE

Social media research comes under the umbrella term ‘internet-mediated research’ (IMR) which is defined by the British Psychological Society as:

“...any research involving the remote acquisition of data from or about human participants using the Internet and its associated technologies.”

(Ethics guidelines for internet-mediated research, BPS, 2021)

This guidance focuses on the gathering of **novel, original data from social media platforms (e.g. user comments, forum posts, metadata)** to be subjected to analysis in order to provide new evidence in relation to a particular research question (adapted from Hewson et al., 2003).

3. RESPONSIBILITIES

Staff and Postgraduate Researchers (PGRs) are responsible for ensuring their on-going awareness of the requirement for ethical review and to utilise the ethical framework of the University. It is also their responsibility to ensure their on-going awareness of and compliance with relevant external governance and legislative requirements, and any standards or requirements set by funders, scientific and professional bodies and any other relevant stakeholders. Furthermore, supervisors of students are responsible for ensuring that they provide guidance on the ethical principles within this document and that they provide appropriate support to their supervisees in submitting and amending applications for ethics review and ensuring they are of a standard necessary to commence research.

Any research project involving social media users/their data will require University ethics review and data collection must not start until the appropriate ethics approval and permissions are in place. This includes taught students conducting social media research where there is no intention to publish data externally beyond course-work submission for examination. For the purposes of ethics review, collecting data from social media is considered to be primary data collection.

4. GUIDANCE ON SOCIAL MEDIA RESEARCH

Key ethical principles

When collecting data from social media, in research terms, the social media users are human participants and the principles of ethical research with human participants apply.

In accordance with the University of Sussex Research Ethics Policy, social media research should, like any other research, ensure that ethical principles are considered and addressed:

- a) Benefits of the research outweigh any risks or harm**
- b) Risk and harm are mitigated as far as practically possible**

The potential risks and harms related to social media research:

- There may be vulnerable individuals embedded in the dataset from whom consent would normally be received via a guardian - e.g., the dataset may include children or adults who lack capacity to consent. A social media user may also be vulnerable at the time of posting and may wish to withdraw their post or they may not fully understand how their information may appear publicly or be used or viewed for research purposes.
- There is a potential to identify or re-identify research participants/social media users in publications or outputs.
- Social media users may wish to withdraw, amend or dissociate themselves from a post or statement, which would be no longer possible to do if it is included within a research dataset. Neither do social media users have a right to reply or to provide context for their statements.
- There may be risks to the individual user through the use of their data when linked to them, for example, disclosure of criminal activity, political views, health information, sexual orientation/activity etc.

- By disclosing information about a social media user, that individual is at risk of distress, humiliation, intimidation (or “trolling”), stalking or “doxxing” (release of sensitive information without consent) or legal repercussions (including transnationally).

c) Conflicts of interest are transparent and managed

Researchers should declare all potential uses and outputs for the research and declare any conflicts of interest in their ethics application (for example, funding sources, data sharing, using with AI).

d) Research participants give informed consent before taking part in research. In the very limited circumstances where this is not possible, for example in social media research, the risks should be considered, justified and managed

Informed consent is one of the core principles of research ethics. To ensure that research is ethical and that respect for the autonomy, privacy and dignity of individuals and communities¹ is upheld, researchers should usually seek informed consent from participants by giving information about the research that allows them to make a meaningful choice about whether or not to take part.

In social media research, however, this is not always possible and “In many cases, a social media user’s data is accessed and analysed without informed consent having first been sought.” (Townsend & Wallace, 2016). This does not mean that the principle of informed consent does not apply to social media research; rather, that it is context-specific and will depend on the extent to which the social media user’s data can be said to be in the public domain, the terms of service of the platform the user posted on, and the nature of the data itself.

In some instances, such as accessing data via an online forum or group, it may be possible to seek informed consent from individuals or from a website owner/moderator who would act as a gatekeeper (if there is consent to share already in place). In other cases, such as projects involving automated mass data-scraping, it will not realistically be possible to do so, and researchers will need to think about providing a public interest argument because there is no consent to use the data for research. There may be a small number of cases for example, public figures and large organisations, where consent could be deemed to be ‘implied’ if social media posts are clearly aimed for public consumption.

The perception of the user as to whether their data is private or in the public domain must be considered. For example, posting on a closed group/forum versus contributing to a hash-tagged public thread on Reddit or X (previously known as Twitter). Even if this is in the public domain, users may not expect their posts to be reproduced or interpreted/analysed in research publications.

You should also confirm whether any special permissions are required to conduct the research such as explicit permission from the platform or gatekeeper access and provide details of these. For example, if you are accessing a closed group or thread, there may be group rules which prohibit scraping data or require group admin approval. An example of research where data was collected from an online forum, was published and later retracted due to privacy concerns –

[Retraction Note: The complex lived experience of schizophrenia diagnosis: a thematic analysis of online forum posts | Current Psychology](#)

¹ [Ethics guidelines for internet-mediated research](#) British Psychological Society, 2021

Researchers cannot rely solely on a user's agreement to a platform's terms of service as equivalence to giving informed consent – there is strong evidence that users do not always read or understand terms of service².

It is not possible to provide a prescriptive list of scenarios where informed consent is required, and those where consent may be implied, as each case will be context-specific. A general rule of thumb is that the use of direct quotes or attributed quotes would typically require consent, but there may be exceptions to this, such as quotes taken from official or public-authority websites or from the content of a published newspaper article; or those made by public figures acting in their public capacity.

Thinking about participants' privacy, confidentiality and anonymity will help researchers work through some of the issues around consent. The researcher should justify in their ethics application as to why their research enquiry might not require informed consent and explain how participants will be protected from harm.

- e) Research participation is voluntary and the risks are considered in the very limited circumstances where this is not possible**
- f) Participants are free to withdraw from the research without consequence, and they are informed of any timeframe for withdrawal. In the very limited circumstances where withdrawal is not possible, this should be clearly explained to participants in advance or risk mitigations put in place**

When informed consent is absent in the research design, researchers must consider that other principles of ethical research, such as participation being voluntary and the ability to withdraw, are also absent.

In terms of data withdrawal, what happens if a person deletes their post, but it remains in a researcher's dataset? Typically, researchers will access social media data via the platform's API service. [An API is an 'Application Programming Interface'](#), a piece of software that researchers use to make data requests that matches their specified search criteria. The API returns a dataset according to the criteria. This dataset is typically refreshed frequently to ensure that it is up to date (e.g., removing deleted posts). Most platforms do not allow the use of third-party automated data collection software such as web scrapers and web crawlers as they circumvent these data checks and disrupt the service.

Where researchers are manually collecting posts from social media, and API safeguards are not in place, care should be taken to anonymise data so that individual posts cannot be traced back to the person.

- g) Safeguarding of those involved in research or research-related activities is considered and appropriate pathways for reporting are identified including researcher safety and wellbeing**

Researchers may encounter safeguarding concerns in their datasets. Both the impact of the content to the researcher and any potential reporting requirements should be considered and addressed in the ethics application.

² [Social Media Research: A Guide to Ethics](#) Townsend & Wallace, 2016

Security sensitive research material is considered to be that which relates to extremism, terrorism, radicalisation and government security measures in respect of these. It may also involve research involving sensitive technologies which may require an export controls licence (in which case, please refer to the [University's Trusted Research webpages](#)).

To ensure that appropriate governance checks are in place, security-sensitive research material (e.g. data that could lead to monitoring by or query from law enforcement or government bodies) must be declared in the ethics application. Research that is considered particularly challenging for UG or PGT level study will require additional sign off by the Executive Dean (or delegate) in advance of seeking ethical approval. Any queries about security-sensitive research and/or the above process should be directed to the Research Governance Office rgoffice@sussex.ac.uk in the first instance.

h) Researcher safety and wellbeing

Personal social media accounts and profiles should never be used for University research/interacting with participants due to the blurring of professional and personal boundaries and the risk to both participants' and researchers' privacy and personal safety. Where it is necessary to have a presence on a social media platform for research purposes, a separate profile should be created for this purpose (using a University e-mail address or a dedicated study webpage etc.).

Staff and students should use University IT equipment or their University of Sussex MS Edge browser where possible. This will ensure that any search terms used are linked to your University profile rather than your personal one and will not follow you to personal accounts or searches. Similarly, personal smart phones are not suitable devices for collecting data as information may be stored or shared without your - or participants' - knowledge. Where this method of data collection is the only option available, this should be explained in the ethics application and a dedicated study phone account/number created for the purpose of the research.

If the data collection is focused on distressing or challenging topics, researchers should also outline how they will manage their own wellbeing by specifying what support they have access to, and how they will check-in with a Supervisor/line manager or University resources, if needed.

Using social media as a tool to aid recruitment of participants

Researchers wishing to advertise a study on social media should upload a standalone study advert that directs those who may be interested to the study information sheet/details or to a dedicated study site/profile created under a University email account/email; OR invites potential participants to contact the researcher directly via a University of Sussex email address. Researchers should not invite potential participants to make contact on the social media platform itself or via a personal email address.

Once posts are created on social media using professional/University accounts, then it is possible to share (or re-post) those official posts via a personal account to reach your personal network (if you have ethical approval to recruit from your personal contacts). It is recommended that you share the post without further edits or additional comments.

We understand that researchers may want to share projects with potential participants from their existing contacts using messaging apps (e.g., WhatsApp). To ensure safety and participant anonymity, do not share study information within messaging groups where everyone can see each other's contact information (e.g., names and phone numbers). If the messaging app allows it, you can share

links to a study via mechanisms that preserve recipient anonymity; for example, in WhatsApp, you can use the broadcast feature to send study links to a list of potential participants without recipients seeing one another during messaging.

Like many personal interactions, researchers who use social media and messaging may encounter harassment. Risks of harassment should be considered in the ethics application, and the researcher should detail how risks are minimised and what support mechanisms are available to address such issues should they arise. There are several online guides available, from organisations as well as social media and messaging apps, that may be useful for helping to ensure safety. Make sure to check out help resources from different apps, as guidance links may change over time:

- [UK Safer Internet Centre Advice](#).
 - [Suzy Lamplugh Trust Personal Safety Advice](#).
 - [Family Lives Bullying on Social Networks](#).
 - [X \(Twitter\) Reporting Abusive Behaviour](#).
 - [WhatsApp Wellbeing Support](#).
 - [Facebook Handling of Bullying, Harassment, or Personal Attacks](#).
 - [Instagram Information on Abuse and Spam](#).
- The latest guidance on this topic is available on the ethics A-Z guide: https://www.sussex.ac.uk/staff/research/governance/apply/ethics_a_z#socialmediarecruitingparticipants
 - More general advice on the University's guidance for social media use and the University policy for staff and social media use is available here: <https://www.sussex.ac.uk/brand/staff/social>

Research protocols adequately address the privacy and confidentiality of participants

Researchers should always seek to protect participants' privacy and confidentiality. Typically, this is achieved through anonymising individuals in research outputs so that quotes/data etc. cannot be linked to the research participant. In social media research this is difficult to achieve. A quick search of a quote via Google or another search engine may identify the individual, their location, related content, or the time and context in which they originally posted and therefore anonymity cannot be assured.

A helpful way of thinking about the nature of social media users' data and its representation in your research is to think of it as "private data on public display" (Nicolas Gold)³. In other words, a person's data may be technically accessible or publicly available, but it still contains private information about them that they may not expect to be 'studied' or included in another medium.

Wherever possible researchers should paraphrase quotes or present outputs in aggregate form to avoid the identification of individuals. Where the use of direct quotes is necessary for the research output, it may be appropriate to contact the participant to seek their consent. A general rule of thumb is that the use of direct quotes would require consent but there may be exceptions to this, such as quotes taken from an official or a public-authority website or from the content of a

³ [Using Twitter Data in Research Guidance for Researchers and Ethics Reviewers](#), Dr Nicolas Gold Department of Computer Science, UCL, 2020 AND [Ethical issues in research using social media - UK Research Integrity Office](#) / <https://youtu.be/RTWWnOV4Gck?si=wvx7EFq1abfQP5X6>

published newspaper article; or those made by public figures obviously acting in their public capacity.

Even where the research output will not identify individuals, the data collected for the analysis constitutes personal data and must be stored securely on University systems in order to assure confidentiality.

Data protection considerations and legislation

Social media research involves the processing of personal data and is subject to the Data Protection Act 2018 and the UK General Data Protection Regulation (UK GDPR).

Personal data is **“Any information relating to a person (a ‘data subject’) who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person.”** Examples of personal data include name, address, date of birth, research participant number, pseudonym, occupation, email, CV, location data, Internet Protocol (IP) address, phone number, a user-name, weblink to a post, the exact wording of the post. *Any* information which can be used to easily identify a participant with an internet search, counts indirectly as personal data.

The University must always have a lawful basis for processing⁴ personal data and, in the context of research activities, our lawful basis is carrying out a task “in the public interest or in the exercise of official authority vested in the controller”, known as ‘public task’ (Article 6(1)(e) of the UK GDPR.) Under our Royal Charter, the purpose of the University is to “advance learning and knowledge by teaching and research to the benefit of the wider community,” meaning research activities are part of our ‘public task’.

Although there are some limited exemptions for research activities, researchers will need to comply with most of the requirements of data protection legislation. The key principles are:

- Personal data should be processed **legally** and in a fair and transparent manner;
- Any personal data must be adequate, relevant and limited to what is necessary for the research;
- **The collection of personal data should be limited to what is adequate, relevant and necessary;**
- Personal data should be accurate and, where necessary, kept up to date;
- Personal data must be kept secure;⁵
- **Personal data should not be kept for longer than is necessary.**

Special category data and criminal offence data

⁴ ‘Processing’ personal data means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

⁵ <https://www.sussex.ac.uk/ogs/policies/information/dpa/research-and-gdpr>

Under data protection legislation the following types of personal data are called 'special categories of personal data'.⁶

- personal data revealing **racial or ethnic origin**;
- personal data revealing **political opinions**;
- personal data revealing **religious or philosophical beliefs**;
- personal data revealing **trade union membership**;
- **genetic data**;
- **biometric data** (where used for identification purposes);
- data concerning **health**;
- data concerning a person's **sex life**; and
- data concerning a person's **sexual orientation**.

Special category data should only be processed for a limited range of purposes and only where necessary. Under the UK GDPR researchers may process special category data for scientific and historical research purposes provided that:

- The processing is in the public interest;
- It is not likely to cause substantial damage or distress to the individual; and
- The processing must not be for the purpose of measures or decisions about a particular person, unless it is necessary for approved medical research.

Social media research projects collecting special category data to meet the research objective must meet the above conditions.

The processing of **criminal offence data** (which includes personal information about criminal allegations, convictions, and proceedings) is also subject to specific conditions and limitations (see <https://www.sussex.ac.uk/ogs/policies/information/dpa/sc-crim-convictions>.)

Mass Data Research

Mass data research involves the use of automated machine learning methods to 'scrape' or collect large sets of data from internet mediated sources; typically, social media. This form of data collection invariably includes data subjects' personal information and will likely include sensitive data, special category data and vulnerable groups within the dataset due to the nature of the auto-collection (even when this data is not specifically sought by researchers). The data is typically collected without the data subject's knowledge or consent.

For the purpose of this guidance, a **large dataset** is defined as that which contains more than 1,000 individuals/data subjects.

If the research will involve the processing of special category or criminal offence data on a large scale, then a Data Protection Impact Assessment (DPIA) will normally be required (see: <https://www.sussex.ac.uk/ogs/policies/information/dpa/dpia>).

⁶ <https://www.sussex.ac.uk/ogs/policies/information/dpa/sc-crim-convictions>

Due to the above requirement, it is recommended that taught student projects do not include the collection and analysis of mass personal special category data or criminal offence data as a project objective. The same recommendation applies to security sensitive data.

In mass data research, it is essential that project design and ethics application consider:

- The inadvertent presence of vulnerable groups (e.g., minors) in the dataset;
- The collection (either targeted or incidental) of special category data;
- The collection (either targeted or incidental) of sensitive data;
- The fact that the data set will exceed 1,000 data subjects;
- [Data minimisation](#) principles, to collect only what is relevant, adequate and necessary

And detail the steps researchers will take to:

- minimise the personal information held;
- pseudonymise the data;
- anonymise the research outputs;
- process personal data safely and securely and in accordance with data protection legislation and University requirements.

Researchers should consider that there may be bots and spam embedded in the dataset which were deployed for malicious intent (hate speech, swaying elections etc.). They may falsely polarise the topic you are trying to research, and these can be increasingly difficult to check the larger your dataset is. This should be addressed in your ethics application, with the data minimisation principles in mind.

Where mass data projects present potential risks relating to information security, data protection, insurance implications and other legal risks, the ethics application will be referred to the University governance teams – IT Services, Legal Team (General Counsel), the Data Protection Team and the Insurance Manager. University governance teams will review the legal and compliance issues in the more complex social media mass data projects and provide specialist advice to the ethics committee to inform the ethics review. The reviewers may also recommend any steps required to mitigate identified risks.

Social Media platforms and their Terms of Service

It is essential that researchers read and follow the terms of service of the social media platform they are collecting data from.

As explained in the [Key ethical principles](#), terms of service are never a proxy for informed consent for participant data, but understanding the nature of the user's contract with the platform will be helpful when negotiating issues around privacy and confidentiality and considering whether an individual's data can be said to be 'on public display.'

Researchers must also be very clear as to the terms under which they may collect the data and present that data in research outputs. Typically, researchers will need to set up a user account to collect platform data and in doing so, agree to the platform's terms of service, which they must abide

by. Reproducing content (such as direct quotes) from the platform may be prohibited. Some platforms require prior approval before their data is used in publications, some platforms may specifically state that their website cannot be used for research purposes (including text/data mining or analysis to generate information which includes but is not limited to patterns, trends and correlations), and most platforms do not allow their content to be used for AI training.

Social media platforms also generally prohibit the automated collection of data (known as 'data-scraping') unless researchers do so via a specific developer agreement that grants them access to software developed/managed by the platform. For example, X mandates that researchers must collect data via the X Application Programming Interface (API). This places limits on the rate of data collection and also enables X to notify researchers when a user has requested the deletion of a post (which should in turn be removed from the research dataset).

When intending to publish analysis of social media data, researchers must check and ensure compliance with the platforms' Terms of Service from the outset. This may include data collection via the platform API and/or seeking permission from the platform itself prior to publishing. The University generally cannot endorse publication of research data which has not been collected in compliance with a platform's Terms of Service. All researchers, (including Taught students) are encouraged to plan for potential publishing during their research design stage and include this in their ethics application, since retrospective approval will not be granted.

It is not possible to explore the various platforms' terms of service within the scope of this guidance. **It is the researchers' responsibility to carefully read and comply with any terms and conditions.**

Third-party datasets

Where researchers wish to use a dataset that has been collected or curated by a third party, the above ethical considerations will still apply. Researchers should demonstrate in their ethics application that the data has been collected ethically, and that its use does not place the researcher or University at legal/other risk.

Please refer to the University guidance on secondary data, available on the [A-Z](#) guidance:

<https://www.sussex.ac.uk/staff/research/documents/ethics-guidance-on-the-use-of-secondary-data-in-research.docx>

Artificial Intelligence (AI) and social media research

The use of AI in research will be outlined in separate University guidance and/or reviewed on a case-by-case basis. Any use of AI software to collect, process, create, analyse, or present social media data should be declared in the ethics application so that appropriate ethical and data protection review can be undertaken. The use of AI will also need to be declared in any research outputs.

Tips on filling in your ethics application

All ethics applications must be submitted on Infonetica Ethics Apply:

<https://apply.ethics.sussex.ac.uk/>

Full guidance on the on-line ethics review application process is available at:

<http://www.sussex.ac.uk/staff/research/governance/apply>.

Additional guidance on the ethics review of studies involving social media data is set out below:

Screening Tool

The application starts with the screening tool. At minimum, social media research will involve human participants and access to/collection of/use of their personal data, so these options should be ticked. You should consider and tick the other options, as applicable to your research.

Screening Tool for Researchers

1.0 Does your project involve any of the following?

☒ Research involving human participants.

☒ Access to, collection or use of personal data or property, including mass data collected online (including from social media platforms).

☐ Access to, collection of or use of non-personal sensitive or confidential data.

☐ Research with the potential to expose any person, whether participating in the research or not, to physical or psychological harm.

☐ Research with the potential to cause a significant negative impact or damage to the environment.

☐ Research involving genetic material and the local or traditional knowledge relating to the genetic material.

☐ Research exploring or involving illegal activities, requiring access to or handling of materials related to illegal activities and/or research that could lead to the disclosure of information that could facilitate illegal activities.

☐ The collection and/or use of material derived from humans.

For the purposes of ethics review, collecting data from social media is considered to be primary data collection:

3.5. Is this research project only using secondary data sources, with no primary data collection?

This does not apply to datasets collected from social media, which will require full ethical review.

☐ Yes

☒ No

Description of the project – rationale and overview of methods

As with any ethics application, the project description should be a clear, easy to read summary that is as jargon free as possible. It should provide an overview of the whole of the research study that readers can understand the first time they read it. The project description should include details of the research that involves human participants/their data, how participants' data will be selected, and a description of the methods used to gather the information. The research question and rationale should include a justification of why you want to use social media data without participant consent.

Key data protection considerations researchers need to address in their ethics application

- **Public task:** Is the research justified as being in the public interest?
- **Proportionality:** Is the processing a targeted and proportionate way of achieving the research purpose? Can you quantitatively show that the scope of data collection is proportionate to the research questions and methods being applied (e.g. range of sources/number of items)? If you exceed this, is it justified? Are you avoiding the storage of data fields (e.g. meta data) that are not being used in the analysis?

- **Confidentiality/Anonymity/Consent:** How is this being assured? Consent will not be informed so it's important to consider 'risk' – e.g., is the data is potentially sensitive/embarrassing or is it about fairly mundane daily activities or opinion? Have you explained how you will avoid or mitigate for the use of verbatim content? Where sensitive or special category data is involved, whilst the UK GDPR provides for the processing of (special category) personal data which are manifestly made public by the data subject, you must still minimise the risk of identification/re-identification and/or the linking of data that can identify the original user.
- **Terms of Service and Developer Agreements:** Have you documented the platform's official means of accessing the data for research purposes - e.g., via an API - and evidenced your compliance with this?

Even where you are not recruiting or interacting with social media users in order to access their data, they are still participants in your research. Consequently, in your application, you will need to provide information about them and how you will assure their privacy, confidentiality and anonymity and mitigate any risks to their wellbeing.

Participants: State how many people/data subjects that you envisage being a part of the research, what characteristics you are looking for and how will they be selected and (if applicable) approached. Collecting data from >1000 individuals should be declared and you should refer to the Mass Data Research section above. It is helpful to describe the 'type' of social media user you anticipate taking part in your research: are they users, sharers, creators, commentators, observers, public personas...?

With automated data collection it is appreciated that the number of data subjects cannot be known in advance but you should still give a rough estimate or specify a data saturation point, as it is important that the data collection is necessary and proportionate (i.e. we shouldn't be collecting more data than we need).

Recruitment: If you are not recruiting/approaching participants you need to explain why this is the case and provide further detail on the selection criteria. Think about the sites you are recruiting from: are they targeting under 18s? If so, there needs to be careful consideration around selection criteria, as data identified or suspected as belonging to children⁷ should not be used without appropriate consent.

You should include a list of your search terms (demographics, followers, hashtags etc) as these are effectively the means by which you are selecting participants.

Data management

Describe the methods and tools you are using to collect and analyse the data. You should also include details of how you will store and de-identify the data and should describe the research outputs (e.g. aggregate tables, paraphrased quotes etc).

All research projects should have a Data Management Plan (DMP) prepared before data collection starts. The library team can advise researchers on preparing a DMP:

<https://www.sussex.ac.uk/library/researchdatamanagement/>

Location: Give details of the social media platforms you are using and how/where you will be accessing the data.

⁷ Children are individuals under the age of 18 years:

<https://researchbriefings.files.parliament.uk/documents/SN07032/SN07032.pdf>

Risk Assessment

The checklist contains 9 questions to help determine whether the project is low risk or high risk. If the answer to all of the questions is 'NO', the project is considered low risk. If the answer to any of the questions is 'YES', then the study is considered high risk. If there is a case for the project to be considered low risk despite a 'YES' answer to one of the questions, this can be made at the end of the section.

In social media research, unless you are seeking informed consent from participants, the answer to checklist question **5.4. Research involving “Deception and/or participation without consent”** will generally be 'YES'. If the answers to all of the other questions are 'NO', then it may be possible to make the case for low risk review. Exceptions to this are where:

- the focus of the research is **personal special category** data; and/or
- the study involves mass data (i.e., the automated data collection of a data set >1000 data subjects);
- where one of the other checklist questions triggers a 'YES' response.

In these instances, the project will be reviewed as a high-risk application.

CONTACTS FOR FURTHER INFORMATION AND GUIDANCE

If you have any queries about social media research and/or this guidance, please contact the Research Ethics, Integrity and Governance Team (rgoffice@sussex.ac.uk) or contact your **Faculty Research Ethics Committee (F-REC)** directly:

Science, Engineering and Medicine Faculty: Science, Engineering and Technology F-REC (SEM F-REC: SET) covering Engineering, Informatics, Life Sciences, Mathematical and Physical Sciences, Psychology

Email - frecsemset@sussex.ac.uk

Faculty of Social Sciences F-REC (SS F-REC) covering Education and Social Work, Global Studies, Law, Politics and Sociology, and students at the Institute of Development Studies (IDS)

Email frec-ss@sussex.ac.uk

Faculty of Media, Arts and Humanities F-REC (MAH F-REC)

Email frec-mah@sussex.ac.uk

University of Sussex Business School F-REC (USBS F-REC)

Email frec-usbs@sussex.ac.uk

Science, Engineering and Medicine Faculty F-REC: Brighton and Sussex Medical School (SEM F-REC:BSMS)

Email SEMFRECBSMS@sussex.ac.uk

Review / Contacts / References	
Guidance title:	Guidance on the ethics of research using human participant data from on-line sources/social media
Date approved:	23/04/2026
Approving body:	Research Ethics and Integrity Sub Committee
Last review date:	23/04/2026
Revision history:	V2
Next review date:	23/04/2029
Related internal policies, procedures, guidance:	University of Sussex Research Ethics Policy University of Sussex Research Ethics Policy Supporting Notes Data Protection Policy Records Management Guidance Use of secondary data in research Research Data Management Policy Freedom of Speech Code of Practice
Document owner:	Research Ethics Integrity and Governance (RIS)
Lead contact / author:	Senior Research Ethics and Integrity Officer Carol Cooley (rgoffice@sussex.ac.uk)