Ethics considerations in Internet-mediated research (IMR)

Internet-mediated research (IMR) is becoming increasingly common. While the core principles of ethical practice remain the same, features of online environments can make certain issues more prominent than in traditional research. As technologies evolve, so do the ethical challenges associated with IMR.

Ethical considerations can arise at every stage of a project—from planning and design through data collection, analysis, and dissemination. This guidance aims to support researchers in identifying and reflecting on these issues early in the design process.

The guidance draws on:

- British Psychological Society's Ethics guidelines for Internet-mediated research,
- Janet Salmons' book Doing Qualitative Research Online (Sage, 2022),
- Ethical guidelines in internet research produced by the Association of Internet Researchers,
- Internet Research Ethics (Stanford Encyclopedia of Philosophy)
- LSE guidance on using data from internet and social media research
- British Sociological Association's guidelines on Researching online forums,
- and
- The Medicines and Healthcare Products Regulatory Agency's guidance on eConsent.

Rather than applying fixed rules, ethical decision-making in IMR requires a proportionate, context-sensitive assessment of privacy, consent, identifiability, risk, and potential harm.

All IMR research needs to receive a favourable approval from the relevant College Research Ethics Committee (CREC) before it begins**.**

What is internet-mediated research

Internet-mediated research (IMR) broadly refers to any research that collects data from or about human participants remotely. It can use qualitative, quantitative, or mixed-methods designs. IMR may be *reactive*, where participants actively engage with materials or a researcher (e.g., online surveys, online interviews, or online focus groups), or *non-reactive*, where data are gathered unobtrusively from secondary sources not created for research purposes (e.g., analysing social media posts or observing chatrooms, data or web scraping). Just as researchers must follow UWE's

ethics policies and review processes, they are also required to honour the rules of the online platform they use. This includes respecting the site's specific norms and community expectations, as well as adhering to any legal obligations outlined in its terms of service.

An essential consideration to keep in mind in IMR is that collecting data without physical co-presence has implications for ensuring valid consent, enabling withdrawal, and protecting participants.

These features mean that the ethical acceptability of IMR depends not only on the method used, but also on the context in which data are generated, accessed, and interpreted.

While IMR can have a wider reach, since it bridges geographical distances, researchers must recognise that cultural norms around privacy differ and that online methods may exclude individuals with limited internet access or digital literacy. This can further marginalise groups already underrepresented in research and it could also impact the generalizability of research findings.

MAIN ISSUES

Public-private distinction

In non-reactive IMR, the distinction between what is "public" and "private" online is not clear-cut, especially because people's own expectations of privacy differ. Individuals may behave in publicly accessible spaces yet still assume their activity is private, some may lack the skills to adjust privacy settings, or fully understand the implications engagement with social media has on their privacy. This is exacerbated when people use avatars and have different online personas that can often be traced to their off-line identity. Matters become even more complicated when privacy settings change, making previously private content public, or when deleted material remains stored in public archives. IMR also enables the collection of large datasets where the public or private status of the data may be uncertain.

Researchers need to think beyond the simple public/private distinctions and instead use context-sensitive approaches, which consider user expectations, potential harms, and community norms. Ethical use of (big) data requires more than meeting legal or technical requirements—it calls for careful attention to power imbalances, marginalisation, and the lived experiences of the people whose data are being analysed.

Therefore, technical accessibility does not automatically translate to data being ethically available.

Confidentiality and security of online data

The risk of breaching confidentiality is generally higher in IMR than in offline research, both during data collection and when sharing findings. Because researchers do not fully

control the online systems they use, it may be impossible to guarantee complete confidentiality of participants' personal information.

Truly anonymous data are rare. Even an IP address stored with a survey response can be linked to an individual. Datasets that appear anonymous may be re-identified when combined with other publicly available information or traced back to their original source.

Researchers must therefore consider how their data collection, processing, reporting, and interpretation might increase privacy risks or expose participants to harm. The principle of proportionality applies: when confidentiality risks are elevated, researchers should minimise those risks as far as possible and ensure participants are clearly informed of any residual risk to confidentiality.

Participant anonymity

Anonymity is closely tied to confidentiality, as anonymising data helps protect confidentiality. In IMR, anonymisation must cover not only obvious personal details like names or addresses but also information such as IP and email addresses. The software you use for data collection has different settings where you can adjust anonymisation options. If using Qualtrics, when in survey view, go to Survey Options. Scroll down to "Survey Termination" and check off "Anonymize response." This will prevent Qualtrics from collecting any identifiable information including contact information, IP addresses, etc.

When researchers need to link responses over time or allow participants to withdraw their data later, it is better to ask participants to create a memorable code. Such memorable codes need some substantial thought as they themselves should not be identifying. If identifying details like emails must be collected (e.g. for a prize draw), they should be stored separately from the study data.

Researchers should remove uncommon hashtags that could reveal the original source, and treat online pseudonyms with the same respect as real names. They should also avoid naming websites or forums if doing so could compromise anonymity or negatively affect the online community; when uncertain, consulting moderators may help.

Direct quotes from forums or social media should not be reproduced verbatim, as they can be easily traced back to the person who posted the comment. Instead, they should be paraphrased while preserving the original meaning.

Anonymisation may be less likely to be needed/required where: (a) the creator is a public figure acting in their public capacity; (b) the material is genuinely non-sensitive and low risk; or (c) the material is such that its discussion is justified in the public interest. Even then, researchers should still consider whether reproducing or linking the material could create avoidable harm.

Consent

A strong informed-consent process is essential for maintaining the integrity of a study and ensuring participants understand any risks. Valid consent is required whenever online data cannot reasonably be treated as publicly available or when using it without disclosure cannot be justified by scientific value (as outlined in the Code of Human Research Ethics, 2021).

In reactive forms of IMR—such as interviews, focus groups, or surveys—consent is always required. The same applies to participant-observation studies where researchers post in forums, blogs, or social networks and then analyse the responses. Extra caution is required when obtaining informed consent from groups that may be vulnerable to coercion. Researchers will often need to use additional procedures—sometimes offline—to secure permission from parents or guardians before involving minors or other vulnerable individuals in online studies.

When researchers simply observe online interactions, undertake document-analysis studies, analyse posts from a linguistics point of view (e.g. metaphors used by influencers), or collect user-generated posts or images, and do not interact with the creator of the content, the need for informed consent is less clearly defined.

Researchers should keep in mind that even though social-media posts may be publicly visible, they are still created by individuals, and using this material in research can unintentionally draw attention to people who were previously not widely noticed. This increased visibility can make them identifiable and expose them to trolling or other forms of harm. The danger is even greater in politically sensitive contexts, where being linked to certain content could have serious repercussions, especially when platforms do not reliably safeguard users' privacy.

Ethical issues must be considered, and if there is any risk of re-identification, researchers should generally seek consent by contacting the group/site moderator, author, or platform. This can be treated as collective consent—for example, through a site notice, platform permission, or an opt-out message explaining the research. Such notices may need to be pinned or reposted to reach the widest audience. At the same time, researchers should consider whether announcing their presence on a forum could alter how people interact there, making the study ethically problematic. For example, informing a marginalised community about data collection could disrupt the space. When taking this approach, it is considered best practice to first check in with the site's administrators or moderators and to act in line with the community's norms, while upholding ethical principles such as respecting members and preventing harm to the group. Sometimes the researcher may need to proceed without seeking consent, but only after consulting with the relevant CREC to ensure the community is protected from identification and any potential harm or exploitation linked to the research.

Unless expected differently by your specific subject area, UWE guidance is that in observational IMR, consent decisions should be based on a proportionate assessment

of privacy expectations, identifiability, sensitivity, vulnerability, feasibility of consent, and risk of harm.

When informed consent cannot be obtained, it is best practice to avoid using direct quotations—either by focusing on the themes in the data or by paraphrasing any quotes instead. Even if consent cannot be obtained, or is not needed, the study still needs to receive a favourable opinion from the relevant CREC.

Related to issues of consent is also data sharing. Researchers should not assume that data collected through internet-mediated research can automatically be shared for secondary use; particularly if collected through internet data scraping. If explicit consent for data sharing has not been obtained, then consideration should be given to sharing aggregated data, paraphrased extracts, or plausible synthetic examples, as opposed to making the harvested data available. In some cases, it may be unethical or inappropriate to share the underlying data at all.

Ways of obtaining consent

In IMR, participants will normally give eConsent. eConsent refers to the use of electronic media (text, video, audio, websites, etc.) to provide study information and collect consent via devices such as phones, tablets, or computers.

Participants sign consent electronically. Electronic signatures can include signatures that are:

- Tick box plus declarations
- Typewritten
- Scanned
- An electronic representation of a handwritten signature
- A unique representation of characters
- A digital representation of characteristics, for example, fingerprint or retina scan
- A signature created by cryptographic means

Electronic signatures fall into three categories—*simple, advanced*, and *qualified*—and the appropriate type depends on how much trust and verification a study requires.

Simple electronic signatures: Basic forms such as typed names, tick boxes, finger- or stylus-drawn signatures, unique character strings, or fingerprint scans.

Advanced electronic signatures: Uniquely linked to the signer, able to verify their identity, under their sole control, and able to show if the signed data has been altered.

Qualified electronic signatures: A specific type of advanced signature created with a certified device and backed by a qualified electronic-signature certificate, offering the highest level of assurance.

For research that does not involve an Investigational Medicinal Product, *simple* electronic signatures are usually sufficient.

Researchers must be confident that the signer is who they claim to be, that the consent form has not been altered, that the timing of the signature is reliable, and that this can be demonstrated if needed.

For online surveys, researchers can capture consent through the online survey form which they use to collect data. Participants need to give consent to be able to progress to the survey questions.

For online interviews and focus groups discussions the following forms of consent are available:

1. Verbal consent at the beginning of the interview. This must be done using a UWE approved platform such as Microsoft Teams. The verbal consent is recorded at the beginning of the session. Once the verbal consent has been obtained, the recording is stopped and file is stored in a separate location to the file of the interview/focus group which is recorded separately.

2. Participants can provide electronic consent via email if using an organisational email account where the email account is verifiable.

3. Researchers can capture consent via an online platform such as Online surveys, MS forms (providing it being used as part of your UWE Microsoft Office 365 access), or Qualtrics.

4. Participants can download the consent form, sign it, scan it, and send it back, or use their scanned signature to sign the form.

Multiple methods of obtaining consent may be used in a study. The methods used to obtain consent should be proportionate to the level of risk in the study and the degree of assurance required about the participant's identity.

Any personally identifiable data (e.g. name or email address) must be saved in a folder separate to research data.

Deception and debriefing

In some research designs, it may be necessary to withhold certain details or obscure the true research question before collecting data, for example to prevent influencing participants' behaviour and compromising the study's validity. This can however raise additional ethical considerations. In offline studies, these concerns are usually addressed by providing a debrief at the end, explaining the real purpose of the research and reassuring participants. In IMR this is more complex since there are instances where people discontinue their participation half-way through the study and may not see debriefing information. Research Ethics Committees should consider the scientific justification for withholding information or using deception, while also weighing the

possibility that participants might leave the study before receiving the disclosure and debrief and assess any potential harm that could result from this.

Withdrawal rights

Part of informed consent is that participants have the right to withdraw, and participants should be informed whether this is possible, whether any time periods apply, and how to withdraw. In some types of IMR, withdrawal rights can be more complex. In online survey for instance, participants may withdraw without the researcher noticing, sometimes after part or all of their data has already been submitted. This is common in online surveys or experiments, such as when someone simply closes their browser. In these cases, it may be unclear whether they intended to withdraw consent for the use of previously collected data, and using such partial data could violate their withdrawal rights.

Therefore, IMR must anticipate these issues and ensure withdrawal procedures are clearly communicated and as robust as possible.

Similarly, in unobtrusive data collection it is not always clear what happens if the original post was removed before the study ends. Once data are extracted, researchers rarely revisit the source. If the material was public and handled ethically, withdrawal is usually unnecessary unless it involved illegal content or breached platform rules. In such cases, proportionality applies: potential risks must be balanced against scientific value, data quality, and practical constraints. Researchers should, however, make sure that they have taken the necessary measures to protect people's anonymity and confidentiality (e.g. not using verbatim quotes).

Where researchers rely on recently posted material, they should consider whether a short delay before data capture is appropriate, this is particularly so where individuals may still edit or remove posts and where withdrawal expectations may reasonably apply.

Implications of IMR for scientific value

Emerging new methods can introduce new ethical challenges, which should be considered when exploring IMR. In particular, the growing use of AI and machine-learning techniques—enabled by large online datasets such as social media—can allow researchers to infer or predict information about individuals or groups that goes beyond what a human observer could readily see (e.g. individuals' likely personal characteristics, risky behaviour, existing medical conditions). Sharing information about individuals or groups that has been generated through algorithmic inference, and is not necessarily accurate, can cause harm, especially when those inferences may be unreliable or when the groups involved already face social stigma. Additionally, researchers should think of implications that automated decision-making processes have on GDPR and should conduct a Data Protection Impact Assessment (DPIA).

The greater distance between researchers and participants in IMR reduces control over key aspects of the study—such as who is taking part, the conditions in which they respond, their reactions during the process, and differences caused by their devices or software. These inconsistencies can threaten data quality, and repeat submissions can further undermine validity. If using Qualtrics, Qualtrics offers useful advice on Security Survey Options, including on how to address potential repeat submission and advice on fraud submission. Researchers must therefore assess how much control a study requires before choosing an IMR approach.

If methodological limitations are substantial, then the ethical justification for using IMR is weakened, as participant burdens and risks may not be offset by sufficient scientific merit.

Online questionnaires and interviews also face the risk of bots, other AI systems, or non-eligible or non-genuine individuals posing as real participants. Please refer to UWE guidance on non-genuine participation in online research.

Social media spaces are not neutral. Platforms include features like discussion walls, threads, and comment areas, which are all shaped by the commercial companies that run them. Interactions are monitored, and posts are owned by the platform. Some sites also provide explicit guidelines or paid options for accessing live or archived content as research data.

Potential harm for participants and researchers

In addition to the risk of breaching participants' confidentiality discussed above, researcher should also keep other aspects of IMR in mind. Because researchers cannot easily monitor participants' reactions online, there is greater potential for harm, especially in studies involving sensitive topics. Alternative safeguards may therefore be needed.

Similarly, in reactive IMR, using third-party transcription services or AI tools may expose sensitive audio data and should be avoided. Real-time online interviews must be conducted on secure, privacy-compliant platforms, using protected meeting rooms and secure internet connections rather than public networks.

Researchers must also consider their own wellbeing. IMR can expose them to large amounts of sensitive or distressing material, making emotional strain more likely than in some offline research. Using online communities for recruitment or dissemination may attract unwanted or upsetting messages, including unsolicited requests for help. The high visibility of online research can also increase reputational risks for researchers and their institutions.

Risk assessments should therefore address potential harms to researchers as well as participants. Researchers must also avoid compromising their legal or professional standing - for example, by accessing websites linked to illegal activities without proper authorisation.

IMR may not be appropriate for a study if potential harms to either participants or researchers cannot be adequately managed or adequately mitigated.

Questions for reflection

Ethical research practice can be guided by five key questions and researchers should ask themselves the following questions:

1. **Protecting participants:** Have all potential risks been identified, minimised, and clearly communicated, and are institutional, disciplinary, and legal requirements being met?
2. **Securing informed consent:** Have participants been fully informed about the study and their voluntary involvement when necessary/required for the project in question?
3. **Respecting the research site:** Have all necessary permissions for accessing sites, communities, documents, or archives been obtained?
4. **Safeguarding data and identities:** Has the study been designed to protect privacy, anonymity, and confidentiality—including for bystanders—and can data be securely stored, analysed, and reported without exposing individuals?
5. Is IMR methodologically appropriate for the aims of the study, and do its research benefits justify any residual ethical risks?

It is always important to weigh the potential benefits of a study against its risks. For research involving sensitive subjects or vulnerable groups, or when participants' age cannot be verified, it may be determined that an internet-based approach is not ethically suitable.