Chapter 7

Confidentiality

Contents

7.1	Introduction.....	150
7.1.1	The Privacy of Facebook.....	150
7.2	Defining Confidentiality.....	153
7.2.1	What Is Confidentiality?.....	153
7.2.2	Confidentiality and Trust.....	154
7.3	Securing Confidentiality.....	154
7.3.1	Informed Consent.....	154
7.3.2	Difficulties with Informed Consent.....	155
7.4	Anonymity.....	156
7.4.1	Anonymity Versus Confidentiality.....	156
7.4.2	Managing Anonymity.....	156
7.4.3	How to Secure Anonymity?.....	158
7.4.4	Is Complete Anonymization Possible?.....	159
7.4.5	Anonymization in Qualitative Research.....	160
7.5	Breaches of Confidentiality.....	162
7.5.1	What Constitutes a Breach of Confidentiality?.....	162
7.5.2	Culpable Breach of Confidentiality.....	162
7.5.3	Justifiable Breach of Confidentiality.....	163
7.5.4	Enforced Breach of Confidentiality.....	165
7.5.5	Waiver of Confidentiality.....	167
7.6	Conclusions.....	167
7.6.1	Summary.....	167
7.6.2	Discussion.....	168
	Case Study: Too Much Information? A Case Study on Maintaining Confidentiality.....	168
	Assignment.....	170
	Suggested Reading.....	170
	References.....	170

Electronic Supplementary Material: The online version of this chapter (https://doi.org/10.1007/978-3-030-48415-6_7) contains supplementary material, which is available to authorized users.

After Reading This Chapter, You Will:

- Know what confidentiality entails and why it is important in research
- Become familiar with techniques of securing confidentiality
- Understand the difference between anonymity and confidentiality
- Be able to identify breaches of confidentiality

Keywords Anonymity · Autonomy · Blind protocols · Confidentiality · Data leakage · Deductive disclosure · Dignity · Hacking · Identifiers · Informed consent · k-anonymity · Personal data · Privacy (attacks) · Proxy consent · Re-identification · Research data management plan [RDMP] · Self-disclosure · Trust

7.1 Introduction

7.1.1 *The Privacy of Facebook*

Since the launch of Facebook as a (commercial) social media platform, its potential as a treasure trove of data on the dynamics of social networks and both online and offline behavior was quickly recognized by sociologists. In 2006, just a few years after Facebook entered the public sphere, a group of researchers downloaded the publicly available data for an entire cohort of some 1700 freshmen students at an undisclosed US college. The data haul included demographic, relational, and cultural information for each individual, and the interested sociologists intended to use it in generating multiple future research projects.

The researchers had requested and obtained permission to utilize the data for research purposes from Facebook, the college involved, as well as the college's Institutional Review Board (IRB). Notably, they did not seek consent from the individual users, although steps were taken to ensure that the identity and privacy of the students remained protected, including the removal of identifying information (such as the names of the students). Also, they demanded that other researchers who wished to use the data for secondary analysis would sign a 'terms and conditions for use' agreement that prohibited any attempts to re-identify the subjects. Convinced that this would ensure confidentiality, the data set was released in 2008 as the first installment of a data sharing project purported to run until 2011.

However, just four days after the data's release, Fred Stutzman, a PhD student, questioned the non-identifiability of the data, writing: 'A friend network can be thought of as a fingerprint; it is likely that no two networks will be exactly similar, meaning individuals may be able to be identified in the dataset post-hoc' (quoted in Zimmer 2010, p. 316). Soon thereafter, it was established that the source of the data was very likely Harvard College, and although no individual subjects were identified at that point, the dataset was taken offline as a precaution.

In a discussion of this case, Zimmer (2010) observed that the researchers who initialized the project made two assumptions. Firstly, they believed that even if the data set were 'cracked' (allowing individual subjects to be identified), the privacy of

the subjects would not be violated because the data was already public in the first place. Secondly, they assumed the research ethics implications had been sufficiently observed by consulting the college’s IRB and taking steps to anonymize the data.

Addressing both arguments, Zimmer argued that extensive collection of personal data over a long period of time, made publicly available for the purpose of social networking only, by subjects whose consent was neither sought nor received *does* constitute a violation of their privacy (Box 7.1). Additionally, Zimmer found it to be a breach of research ethics because subjects were not provided access to view the data to correct for errors or request the removal of unwanted information (for further discussion of this case, see Zimmer 2010) (Fig. 7.1).

This case raises two important issues. The first being that confidentiality is not merely a matter of shielding research participants’ identities. Confidentiality is about knowing what sort of personal data may be available, to whom, and under which conditions – in essence, it’s about considering the whole picture. It also implies the participant’s right to being informed about the scope of their participation, to being explicitly asked to take part in the research project, and extends to their right to retain (some degree of) control over their own data. Secondly, this case exemplifies how quickly, and recently, our understanding of confidentiality has changed. Not only is it very unlikely that an IRB would approve of the above procedures today, but Facebook and other online social networks have also been increasingly critiqued for their defective privacy policies, of which we have only recently become aware.

In this chapter, we outline confidentiality from a wide lens, and flesh out some of its most salient properties. We will discuss some difficulties with securing confidentiality and examine its relationship to anonymity and informed consent procedures. Finally, we discuss breaches of confidentiality and their consequences.

Box 7.1: What Is Personal Data?

What is defined as ‘personal’ may differ from one person to the next, although there are some obvious instances that perhaps everyone would agree is personal, such as your medical history, sexual orientation, or certain beliefs or opinions. Research policies distinguish between these various categories of ‘personal data.’ The following list, derived in part from the European General Data Protection Regulation, is not exhaustive (Fig. 7.2).

Full name	Sexual orientation
Home address/email address/IP address	Trade union membership
Date and place of birth	Passport/ID/driver’s license number
Mother’s maiden name	Credit card number
Ethnicity/race	Telephone number
Age	Job position
Religious or philosophical beliefs	Biometric records
Political opinions	Criminal record

Fig. 7.1 Can you tell me something personal?

What is personal data?

Name

Address

Localisation

Online Identifier

Health information

Income

Cultural profile

and more

COLLECT
STORE
USE
DATA?

▼

You have to abide by the rules.

Fig. 7.2 What are personal data? (Source: [European Commission](#), data protection)

We restrict our analysis, as we have in all other chapters in this book, to research ethics, and do not cover confidentiality issues within professional relationships, and only briefly touch on the (often severe) judicial components of the issue.

7.2 Defining Confidentiality

7.2.1 What Is Confidentiality?

Any information relating to the private sphere of a person that they wish not be shared with others is considered ‘confidential.’ This information is differentiated from ‘public information,’ which everyone has a right to access. The right of research participants to not disclose certain information and to retain control over their privacy has increasingly been acknowledged inside and outside of academia and has become subject to extensive legislation.

In research ethics, the crucial principle of confidentiality entails an obligation on the part of the researcher to ensure that any use of information obtained from or shared by human subjects respects the *dignity* and *autonomy* of the participant, and does not violate the interests of individuals or communities (see Box 7.2 for clarification of concepts). The right to confidentiality in research is recognized in international bio-ethical guidelines, such as the ‘Helsinki Declaration’ (last updated in 2013), and the European General Data Protection Regulation (GDPR, effective 2018).

In practice, safeguarding confidentiality entails that the researcher observes the following restrictions:

- Research participants remain anonymous by default
- Researchers do not obtain private data unless there is good reason to
- Participants must be briefed on the goal or purpose of the research, its means of investigation, and who has access to the data
- Participants must give active consent, are not coerced to participate, and retain the right to withdraw their cooperation at any moment (even after the study has been completed)
- Participants must be provided with an opportunity to review their data and correct any mistakes they perceive

Box 7.2: ‘Privacy, Autonomy, Confidentiality, Dignity’

Autonomy: the capacity to make uncoerced decisions for oneself.

Privacy: an individual’s sphere of personhood, not open to public inspection.

Confidentiality: private information that a person may not want to disclose.

Dignity: a sense of one’s personal pride or self-respect.

7.2.2 Confidentiality and Trust

Confidentiality pertains to the understanding between the researcher and participant that guarantees sensitive or private information will be handled with the utmost care. Ultimately, confidentiality is rooted in trust.

The participant must trust that the researchers will fulfill their responsibilities and protect the participant's interests. To ensure this happens, an agreement is drawn up in which these duties are specified and communicated to the participant (see Sect. 7.3).

In online and computer-assisted research – a variety that often lacks a face-to-face dimension and perhaps poses a greater privacy threat than traditional research – trust and especially the perception of control over one's own information are key. Addressing the concerns dictates how much and to what degree participants are willing to disclose about themselves (Taddei and Contena 2013).

7.3 Securing Confidentiality

7.3.1 Informed Consent

Perhaps the most important instrument for securing confidentiality is the *informed consent procedure*. It is rooted in the idea that involvement in research should have no detrimental effects on the participants, honor the individual's fundamental rights, and respect relationships, bonds, and promises.

Certain conditions and arrangements have been designed to guarantee safe participation in research. These procedures assume the shape of a contract with a participant who actively and knowingly agrees with the conditions. *Informed consent* typically notifies the participant of the following items:

- Name(s) and affiliation of researcher(s)
- Goal or aim of the research (in comprehensible language)
- Research techniques or procedures to which the participant is subjected
- Risks involved (if any)
- Estimate of time investment
- Agreement on compensation (if any)
- Conditions of confidentiality (anonymization or pseudonymization)
- Storage, usage, and access to data
- Rights of the participant:
 - to withdraw at any moment
 - to review/correct erroneous data (if possible)
 - to receive/be informed about the results (if interested)

- Complaint procedures (including contact details of an independent commission or officer)

Informed consent procedures have become mandatory in social scientific research for qualified researchers, including PhD candidates. Undergraduate students, who do research under the supervision of qualified staff, are generally also required to make use of these procedures (with the responsibility for their proper implementation that of the supervisor). Many of these principles are paralleled by similar procedures in medicine, law, and other professional domains (for further discussion, see Bok 1983, and Israel 2014).

7.3.2 *Difficulties with Informed Consent*

While informed consent thus aims to protect the participant, a few difficulties arise with how we approach it, some of a philosophical nature, others more practical.

One contention is that informed consent is biased towards a particular (Western) view of individuality. Research participants are supposed to be autonomous, well informed, capable subjects who are solely responsible for their own behavior and for knowing with whom formal contracts can be negotiated, and understanding the conditions of their participation.

Not all participants fit into this ideal of autonomous agency. Children (minors), vulnerable communities (for example those who harbor suicidal ideation), or anyone in a dependent relationship who may not be (entirely) free to refuse participation, as well as those who may be unable to fully understand the ‘contract,’ all fall outside of this ideal of autonomous agency.

Furthermore, participants may not always be in the position to appreciate exactly what the conditions of participation entail. This is exacerbated when deception is involved, or when the research design does not allow for the participant to be fully or correctly informed about their role in the study.

Finally, confidentiality procedures warranting subject autonomy favor quantitative research (experimental studies, surveys) that does not require meaningful relationships to be formed with participants. In qualitative research (interviewing, participant observations, etc.) these relationships are pivotal, and formal agreements, such as informed consent procedures, ‘can be problematic in a culture that values relationships over roles and position’ (LaFrance and Bull 2009, p. 145).

Although it is possible to address some of these difficulties in the informed consent agreement between researcher and participant, other issues remain unresolved, especially those regarding qualitative research, to which we return below.

In the final chapter of this book, we review the procedural dimension of confidentiality. There we discuss how to weigh the various risks and benefits, explore how to deal with deception, discuss how to assess the vulnerability of participants and intrusiveness of research, and what to do with ‘chance findings.’

7.4 Anonymity

7.4.1 *Anonymity Versus Confidentiality*

These two concepts, *anonymity* and *confidentiality*, are related, but differ in some important respects. Anonymity can be defined as the degree to which the source of a message can be identified (Scott 1995). It ranges from very high (source is nearly impossible to identify) to none (source is easily identifiable or in fact already identified). Confidentiality, on the other hand, relates to an agreement between the researcher and the participant. The former concerns the initial collection of data, the latter makes promises to not disclose specific personal information.

Seeing as how researchers need to protect the participant from unwanted consequences, anonymity seems a safer guarantee for achieving this goal than confidentiality. A researcher who offers *anonymity* does not record any identifying information. If *confidentiality* is offered, however, identifying information is recorded, but this information will not be disclosed to others.

Does it matter much whether you offer anonymity or confidentiality to your participants? Whelan (2007) demonstrated that research participants are aware of the difference and are equally able to appreciate the different degrees of protection offered under both conditions. This raises the question of whether ‘weaker’ confidentiality agreements could undermine the reliability of research. In a comparative analysis (comparing an anonymous and a confidential condition) of self-reported substance abuse surveys among 15 and 16-year-old Icelandic students, Bjarnason and Adalbjarnardottir (2000) found no evidence that a confidential condition lowered the study’s reliability. Conversely, Lelkes et al. (2012) found that complete anonymity may compromise self-reporting.

Anonymity thus may be a more absolute, though not ‘better,’ criterion than confidentiality in ensuring the participant’s right to privacy. Confidentiality, on the other hand, allows for the creation of a relational dimension that is explicitly left out in anonymity. The importance of relationships in research is a ripe field of study (Box 7.3).

In brief, there can be good reason to offer confidentiality as opposed to anonymity, although anonymity is generally preferred.

7.4.2 *Managing Anonymity*

While anonymity is the norm in research, safeguarding it poses an increasingly prevalent challenge in our digital age. ‘Privacy attacks’ and ‘data leakages’ are rampant and the mechanisms for using public data for participant re-identification have greatly increased (Ramachandran et al. 2012). Netflix’s 2019 true crime documentary ‘Don’t F*ck with Cats’ gives an instructive illustration of how it is possible to identify an anonymous individual from a YouTube video by combining contextual

Box 7.3: Breaking Confidentiality in Good Faith? A Dilemma

Consider the case of a student who did research into ‘workplace inclusion’ at a large governmental institution. The student was commissioned to research the attitudes and experiences of employees with workplace inclusion.

Using a qualitative design, the student interviewed some 20 participants in three different departments in the institution. In accordance with standing institutional policy, the student was not allowed to identify participants on basis of their ethnicity (employees were not ethnicity registered at the institution). However, during the student’s research, she found that ethnicity *did* play a role in how employees experienced feelings of inclusion and exclusion. Some participants felt that ‘the fact that they belonged to an actual or perceived group determined how they were treated by fellow employees and the managers at the institution.’

This result was clearly of importance for the study, yet it conflicted with institutional policy that did not allow the student to identify the ethnic background of the participants. A dilemma arose on how to continue. Should she, or should she not mention ethnicity?

How would you advise the student to proceed? Should the student make use of this information and break confidentiality of the basis that she acts in good faith, or should all mention of ethnicity be removed, in accordance with institutional policy, at the cost of losing vital information?

(Case was communicated to the author. Quotes are altered to prevent identification.)

information in the video (type of electoral receptacles, doorhandles, background noises), publicly available records (street maps, location of shops, etc), and the use of common sense.

Such easy, cheap, and powerful re-identifications not only undermine our faith in anonymization and cause significant harm, they are also difficult to avoid (Ohn 2010). The advantages of digitalization, including increased potential to collect, process, analyze, store, and share data, are countered by new privacy risks, in particular the disclosure of personal data and re-identification. And although GDPR is meant to avoid these risks as much as possible, Rhoens (2019, p. 75) warns how in the age of ‘big data’, power tends to be shifted towards data controllers, reducing consumers’ autonomy, undermining a key element of private law in Europe.

In health-related research there is the ever-present risk that databases get hacked, which are full of sensitive information regarding symptoms, diagnoses, and treatment plans. Longitudinal studies (which follow (groups of) individuals over a long period of time) must allow for an identifying key at least until the study is finished, and thus pose the risk that while the study runs the key is revealed. Online social network analyses that deal with large amounts of data run the risk that the privacy of its users may be infringed upon (as the Facebook example demonstrated). Lastly, as studied by Williams and Pigeot (2017), we should be wary of powerful organizations,

corporations, and governments, who are gathering vats of information about us, further arguing that ‘We have good reasons to fear that this may damage our interests, since their processes of data gathering are so shadowy and unaccountable’ (p. 248).

In an attempt to prepare for privacy attacks and possible data leaks today, research institutions require that anonymization be part of a larger *research data management plan* that invokes policies about data security, data privacy, and data licensing (see Patel 2016) (Box 7.4).

7.4.3 How to Secure Anonymity?

Though this question regards research techniques rather than research ethics, we will have to outline the constraints of this issue before we can discuss the ethical aspects related to it (Fig. 7.3).

The anonymization of data necessitates that *identifiers* (defined below) are changed, either by removing them outright, or by substitution, distortion, generalization, aggregation, or the employment of any number of masking techniques (Ohm 2010).

Direct identifiers, such as name, address, or zip code, can be removed or substituted by a pseudonym or code without loss of information. If substitution by a code is used, a key that allows reidentification may be kept, but as explained above, that key can subsequently become a security risk.

Indirect identifiers, such as information regarding occupation, training, age, or workplace can be aggregated or generalized, but therein lies the risk of information loss (for example, when the value ‘19’ is substituted for the aggregated value ‘between 15-20 years old’).

Box 7.4: Research Data Management Plan (RDMP)

Any RDMP must be compliant with applicable national or international standards and stipulate conditions for the following data-related considerations (pertaining to both new data and amendments of existing projects):

- security, privacy protection, and transparency
- working with sensitive data
- archiving of research data
- storage procedures
- creating metadata
- authorship of data and data use
- verifiability of data
- searchability of data
- data sharing and licensing
- retention period and contact details of the data manager

(Compiled after various university library sources)

Fig. 7.3 Research data management plan. (Source: [Utrecht University](#), research data management support)

Quasi-identifiers make it possible that a participant be identified when disparate information that by itself do not identify a subject are combined to create a clearer picture. For example, in an institution, the combination of age, position, and gender may lead to the identification of the participant if there is only one person with that specific set of characteristics.

In order to anonymize sets of data while retaining as much original information as possible, certain techniques have been developed. One known as *k-anonymity* was specifically designed for quantitative data sets (introduced by Samarati and Sweeney 1998, and since improved, see Amiri et al. 2016).

This technique allows for sensitive data to be recorded but disallows that data may be combined to create quasi-identifiers. Essentially, *k-anonymity* requires that there always be several individuals that match *any* combination of values in the same set of data (see Doming-Ferrer and Torra 2005; Ciriani et al. 2008, for further discussion of *k-anonymity* and Zhou et al. 2008, for a comparison with other anonymization techniques) (Fig. 7.4).

7.4.4 Is Complete Anonymization Possible?

The answer to this question is... probably not. The main reason being that anonymizing techniques, including *k-anonymity*, do not offer fool proof protection against the malicious use of background information, data triangulation, or even just basic web searches (Martin et al. 2007). As a result, 'deductive disclosure' (Tolich 2004) occurs, where certain individual or group traits make them identifiable in research reports.

Fig. 7.4 Anonymization: use of coding, aggregation and redaction (fictitious case). Adapted from Ruth Gerathy, *Anonymisation and Social Research* (2016)

For example, user profiles for many common web-oriented services (such as Facebook or Google profiles) provide a host of background information that make it easy to re-identify research participants in anonymized data sets (Kumpošt and Matyáš 2009). Also, with the aid of publicly available census data that contains records of individual’s birth date, gender, and address, *quasi-identifiers* can be constructed, and anonymized records from smart meter data (Buchmann et al. 2013) or cell phone users (Zang and Bolot 2014) can be used together to re-identify anonymous research participants. Similarly, anonymized online social networks have been de-anonymized with the aid of certain re-identification algorithms (Narayanan and Shmatikow 2009).

The reality is that at this moment, anonymization is not absolute, but a matter of degree (Dawson 2014). A dataset may never be completely safe from intentional attacks, and therefore re-identification of anonymized data presents serious policy and privacy implications (Lubarsky 2017; El Emam et al. 2019).

7.4.5 Anonymization in Qualitative Research

Qualitative research is performed within a diversity of disciplines such as ethnography, anthropological field work, community studies, as well as clinical interviews and (critical) discourse analysis (the study of larger connected sets of textual corpuses). A defining feature of this form of research is that it deals with texts and their non-quantifiable characteristics; the heterogenous and ambiguous structure of language.

When compared to quantitative research, qualitative researchers are confronted with both similar and different challenges regarding anonymization, which we will now explore.

What is *similar* is that qualitative researchers also must consider confidentiality. What personal information are they allowed to make public (with consent from the participant), and what is off limits? When qualitative researchers choose to remove or alter identifiers, they too must worry that background knowledge will allow online sleuths to re-identify (some of) the participants. But masking the identity of

an interviewee or a patient may be even more difficult because of the wealth of self-disclosing information available online.

An additional comparable difficulty that quantitative researchers must also resolve when anonymizing their data, is that even when direct and indirect identifiers are removed, *contextual identifiers* in an individuals' narrative remain. For example, certain (unusual) life events, particular details or circumstances, and specific places are all difficult to anonymize without distorting the text. Despite this difficulty, attempts have been made to create computer programs that facilitate the finding and replacement of personal identifiers (see Kaiser 2009 for a discussion).

What is *different* in qualitative research is that not all researchers share the 'fetish of individualism' (Weinberg 2002). Some insist that research communities are places where people know each other, where ideas and knowledge is shared, and where no-one goes unnoticed. For this reason, they argue, anonymity is virtually unachievable, and the question is whether anonymity is even desirable at all (van den Hoonaard 2003; Walford 2005; Tilley and Woodthorpe 2011). Other researchers have argued to the contrary, and insist that in spite of these objections, anonymity should still prevail as a guarantee against gratuitous and casual identification which does nothing to add to public understanding (Kelly 2009).

Another notable difference with quantitative research is that the 'situatedness' of qualitative data (Thomson et al. 2005) makes secondary use questionable (use of the same data by different researchers). Qualitative data is 'generated through personal interactions with participants, involving a high degree of trust and a duty of care towards participants' data on the part of the researcher' (Irwin 2013, p. 297). Trust and duty cannot be transferred onto unknown researchers just like that.

Finally, Giordano et al. (2007) point out that sometimes participants specifically wish to be heard and do not want to remain anonymous. Depriving them of a voice that provides personal meaning would deny them of (at least a part of) their autonomy. Giordano proposes that participants be offered a choice of disclosing their identity or not.

In light of the discussion above, consider the following study by Wiles et al. (2006). They conducted research about the use of consent procedures among social scientists and health researchers working with vulnerable populations. The participants – experienced researchers who themselves used qualitative methods – were mostly critical of informed consent procedures. Some had little or no experience with consent forms and were put off by the idea of using them. Others refused point blank to sign the forms Wilkes and her colleagues gave them, which they disqualified as an 'overly formalistic and paternalistic enforcement of a biomedical model' (p. 286).

A difficulty was that some of participants were well known in their field and could easily be identified by others in the research community. They refused to give consent that their data be archived. They also insisted that for reason of indefinability, entire sections of the transcripts be deleted. This meant the loss of important findings, while also making secondary analysis impossible.

The 'researching the researchers' study by Wiles et al. (2006) led to the conclusion that in qualitative research, two items of crucial importance cannot be managed

by consent procedures: *trust* of the participant in the research project and their *motivation* to participate in it.

7.5 Breaches of Confidentiality

7.5.1 What Constitutes a Breach of Confidentiality?

A breach of confidentiality means that the obligation of the researcher to ensure their research respects the *dignity* and *autonomy* of the participant is not fulfilled or honored, or that an essential element in the agreement between researcher and participant is broken.

For example, when a promise of anonymity is revoked, then not only is the participant's trust violated, but in the case of any negative consequences the participant may face, the researcher may be liable.

However, *not all breaches are reprehensible*. Some may even be considered justifiable, for example when a higher goal is served. Other breaches may be brought about by a third party and are not a result of the researcher's actions. Or there is the possibility that the breach could simply result from the wish of the participant not to remain anonymous (waiver of confidentiality) (Fig. 7.5).

In the coming section, we discuss examples of these four classifications of breaches in further detail, and identify a number of consequences and possible remedies.

7.5.2 Culpable Breach of Confidentiality

When sensitive, personal, or identifying information from participants is made public without their consent, and it has negative consequences for the participant (or the community), the researcher can be held responsible if they could have prevented this from happening.

Suppose a researcher interviews a number of employees from several different companies about their job satisfaction. The participants are guaranteed complete anonymity. At some point in time a report on the findings is published. Now consider that a supervisor at one of the participating companies reads the report and is able to ascertain a certain participant as one of their employees, based on a number of characteristics. Since this employee has been found to be critical of the

Fig. 7.5 Classification of breaches of confidentiality

Culpable breach of confidentiality	Justifiable breach of confidentiality
Waiver of confidentiality	Enforced breach of confidentiality

organization, the supervisor decides not to promote them. Now, the question can be asked: is the researcher responsible or even liable for the damage?

The answer depends on whether the researcher has *done enough* to secure the anonymity they guaranteed. For example, if only the participants' names were anonymized, but certain unique characteristics remained (gender of the participant, the position in the organization, age group), allowing for easy re-identification, then the researcher might indeed be liable. But if the researcher has done everything that reasonably could be expected from them, and the supervisor deduced the identity of the employee by chance, the breach of confidentiality could be considered merely lamentable, not culpable.

In 2015, the journal *Expert Systems and Applications* published a paper that used several sentences taken from the logged-in section of a website called 'PatientsLikeMe'. One particular quote from an HIV-positive user on the site contained specific dates and types of infections the patient had experienced. This led to a complaint to the editors of the journal that 'a search within PatientsLikeMe for this string [of information], or fragments of it, would quickly identify this patient.' The editors of *Expert Systems and Applications* accepted the validity of this complaint and withdrew the paper. The authors were requested to delete the incriminating quotations and when completed, the paper was later republished (case taken from 'Retraction Watch,' September 2016).

7.5.3 Justifiable Breach of Confidentiality

There can be good reason to disclose sensitive or personal information without consent if circumstances force the researcher's hand. This can be found, for example, if a third party could find themselves in immediate or future risk were certain information to remain unknown to them (Box 7.5).

The well-known 1974 'Tarasoff Case' may be taken as the classic example of a justifiable breach of confidentiality, even though it is more of an instance of professional ethics rather than research ethics (see Bersoff, 2014 for a discussion). In the 'Tarasoff Case,' a patient confided in a mental health professional his intentions to kill someone. The intended victim was not alerted and indeed, was later murdered by this patient. When the case came before a court of law, it was ruled that client-therapist confidentiality *should* have been breached because of a higher duty, the protection of the intended victim (see Herbert 2002 for a discussion of the case).

In social science research, analogous situations may present themselves, even though they are rarely as extreme as the Tarasoff Case (see Duncan et al. 2015). For example, Jonathan Kotch (2000) discussed some of the challenges researchers encountered in studying the complex matter of longitudinal child maltreatment research, which led to serious deliberations regarding breaching confidentiality.

The researchers were interested in the behavior of mothers, but in the process they not only collected confidential information about children, but also from them. Would this make these children automatically research participants? And if

Box 7.5: Breaching Confidentiality or Not? A Dilemma

George is a psychologist who is interested in high-risk sexual behavior among adolescents. He speaks with 25 participants, one of whom is Martin, an 18-year-old male who has agreed to be interviewed on the provision of complete confidentiality. During the interview, Martin reveals that he has been diagnosed with HIV and has not informed his partner, even though they have regular unprotected sexual intercourse.

George is worried that he is obliged to breach confidentiality and disclose this information to Martin's partner. For guidance, he consults the *Ethical Principles of the Psychological Association*. It states that confidential information can be disclosed (without the consent of the individual) 'when mandated or permitted by law for a valid purpose such as to protect the client, patient, psychologist, or others from harm' (quoted in Behnke 2014).

The laws in George's country aren't very clear about this issue, though. HIV is a contagious disease but doesn't pose an imminent risk of death, though being infected could be deemed considerable harm.

Are there sufficient grounds for George to breach confidentiality? Argue from one of the following positions:

1. George should inform Martin's partner and does not have to inform Martin about this breach of confidentiality because the partner may be in immediate danger.
2. George should inform Martin's partner but also inform Martin about this breach of confidentiality.
3. George should urge Martin to inform his partner but does not have to interfere himself.
4. George should not interfere in any way as he is bound by confidentiality and the responsibility is Martin's alone.

(Case adapted after Hook and Cleveland 1999)

so, under which conditions could they be considered 'participant' in the research? Logically, parents would have to give consent on behalf of their children (this is called 'proxy consent'), on the presumption that they act in the best interest of their children. But that may not be likely in the case here, given that the research was on child abuse and neglect.

Confidentiality issues were further complicated when suspicion of child abuse arose. Under US law, anyone who suspects maltreatment of a child is legally required to report it. Under these circumstances, is longitudinal research on child maltreatment possible at all?

If the answer is yes, then whose interests prevail: those of the mother, the child, or the researcher? Kotch (2000) argues that all three must be factored in when moving forward with a research project, but they carry different weights. Kotch contents

that a child's participation on a basis of 'proxy consent' is ethical as long as the benefits (child welfare, possible beneficial research outcomes) outweigh the risks (harm done to the child). If the child's welfare is at stake, confidentiality may justifiably be breached, but this must be considered very carefully, and weighed against the consequences. This is because the consequences can be substantial, both for the mother (social, personal, and even job-related repercussions) as well as the child (embarrassment, emotional distress).

In order to sensibly handle confidentiality, a special 'blind protocol' was designed for this case, that allowed the mother to respond in writing to sensitive questions that might lead to a suspicion of abuse or neglect, without the interviewer being aware of the answer. Only the principal researchers (PI) would be allowed to review this sensitive material and only they could decide (after careful deliberation) that a case needed to be reported (they eventually did so in five cases out of 442, one of which was confirmed).

7.5.4 Enforced Breach of Confidentiality

There are only a few circumstances that could force a scientist to breach confidentiality. One of those is the enforcement of state regulations. Rik Scarce was a PhD student at Washington State University doing research on an environmental movement in the United States. In his research, he conducted interviews with environmental activists in this movement. Even before his dissertation was published, one of his interviewees attracted the interest of the police. They requested that Scarce appear at the campus police station, where he was interviewed. When he refused to answer certain questions about his research participants and did not allow the police access to his data on grounds of confidentiality, he was subpoenaed to appear before a grand jury. In his testimony, he declared the following: 'Your question calls for information that I have only by virtue of a confidential disclosure given to me in the course of my research activities. I cannot answer the question without actually breaching a confidential communication. Consequently, I decline to answer the question under my ethical obligations as a member of the American Sociological Association [...]' (Scarce 1995, p. 95). This defense was not accepted. Confidentiality simply did not matter to the court, Scarce later wrote (1995, p. 102). He was found in contempt of court and held in custody for over five months.

Although no general conclusions regarding what to do in such cases may be drawn from this case, because laws with respect to liability differ in every country, students should be advised to ensure that their research proposals are in accordance with university policy. In case of doubt, they may want to consult their IRB. See Box 7.6 for further considerations.

Box 7.6: The Russel Ogden Case

In 1994, Russel Ogden, a Canadian MA student in criminology at Simon Fraser University (SFU), completed a controversial study on assisted suicide and euthanasia of persons with AIDS, which was illegal at the time, and attracted a high amount of media attention.

Shortly after the defense of his thesis, based on interviews with people involved in this activity, Ogden was subpoenaed by the Vancouver Regional Coroner to reveal his sources. Ogden refused on grounds that he had promised confidentiality, and that he had acted in accordance with universities policy.

It is noteworthy that Ogden had actively sought approval from the university's independent IRB, noting that anonymity and confidentiality would be assured with each participant. He also informed his participants in a consent letter that the 'proposed research project involves data about illegal behavior,' and that participants would not be required to give information regarding their identity. Finally, Ogden sought advice from the university's IRB about what to do in the unlikely event that the Coroner's Office requested cooperation. He was informed that there was 'no statutory obligation to report criminal activity,' and thus accepted full responsibility for any decision he would make (quoted in Blomley and Davis 1998).

When he was subpoenaed, his former university refused to offer assistance, on grounds that 'in cases where it can be foreseen that the researcher may not legally be in a position to ensure confidentiality to their subjects, these researchers must be required to prove only limited confidentiality' (quoted in Lowman and Palys 2000, p. 4). They offered limited financial support only, on compassionate grounds.

Ogden felt abandoned and believed that SFU was unwilling to protect his academic freedom as a researcher. Therefore, after successfully having defended his case before the Court, he filed a lawsuit against the university, claiming they had a contractual obligation to support his ethical stand and to reimburse his legal fees.

Though Ogden lost that case, following Bloomley and Davis' 1998 review of it, the university belatedly accepted responsibility and reimbursed his legal fees and lost wages and send him a letter of apology, promising to assist researchers in the future who may find themselves in the position of having to challenge a subpoena (see Lowman and Palys 2000, for a discussion of the case). Ogden later became a world leader in suicide research, but his work remained controversial.

7.5.5 Waiver of Confidentiality

Finally, we consider cases where participants themselves wish to be identified or wish to waive their right to confidentiality. Technically these would not be breaches of confidentiality, but rather waivers of confidentiality agreements.

The number of cases in which participants waive confidentiality and/or in which IRBs agree to such a design are uncommon. In certain types of research, however, waivers of confidentiality are the rule rather than the exception. In Participatory Action Research (PAR), participants agree to be ‘collaborators’ of the researchers, not ‘research subjects.’ They will not merely be ‘interviewees’ or ‘respondents’ of the researcher, but actively engaged in the research process itself, defining together with the researcher the research question and research set up. This means to a degree, the roles of researcher and participant roles blur. And as much as this is the case, there is good reason to give ‘special concerns regarding the need to protect confidentiality’ say Khanlou and Peter (2005, p. 2338), although that does not necessarily imply lifting it. It means that participants themselves decide how they be involved and define their involvement.

There may be another reason for participants to give confidentiality a second thought. Vainio (2013, p. 689) examined an example in which a researcher conducted a study of an organization, and the individual who developed the organization insisted they be mentioned by name in the report (in the hopes of profiting from it). Here, waiving confidentiality borders on a conflict of interest (see Chap. 8).

7.6 Conclusions

7.6.1 Summary

Confidentiality stands as a core tenant of scientific research ethics. Few issues matter more than allowing the participant control over which information they wish to share. The most important procedure in this aspect of research is to secure the *informed consent form*, which formalizes a confidentiality agreement between the researcher and participant. Also, various data points, or *identifiers*, that allow for the re-identification of participants, are important for researchers to understand, as are the techniques to anonymize data, though none offer waterproof guarantee against re-identification. Furthermore, we noted that anonymization in qualitative and quantitative research differs greatly. Finally, *breaches of confidentiality* were discussed, including which forms are justifiable (or excusable), and which are not.

7.6.2 Discussion

Two obstacles regarding confidentiality remain. The first regards the availability of information, and the growing capacity to combine information on a large-scale is making it increasingly difficult to guarantee anonymity. The second is that data protection regulations are still evolving, and the way these regulations coalesce may significantly influence future research agendas.

These two issues – protection of participants' privacy and their autonomy, and evolving data protection regulation – comprise an underlying dilemma: how do you ensure academic freedom while at the same time making sure that everything is done (morally and legally) to protect confidentiality?

Case Study: Too Much Information? A Case Study on Maintaining Confidentiality

The case outlined below highlights some of the difficulties of maintaining scientific standards while simultaneously offering confidentiality, specifically when researching a highly sensitive subject. The following details derive from a group of master's students and their supervisor who were engaged in a field study examining the risks related to sexual and reproductive health (SRH) in a country considered conservative in several respects. Notably in this country, it is a cultural taboo to speak publicly about SRH issues, and accessibility to SRH services remain quite limited as well.

State officials in this country admit that a lack of knowledge on SRH can result in risky sexual behavior and unintended pregnancies, and that these in turn contribute to high rates of sexually transmitted diseases and increased maternal mortality due to (illegal) abortions. While it seems clear that this would justify setting up SRH facilities, a clear government policy on the matter was still lacking, and the emphasis was on countering maternal morality rather than communicating knowledge.

However, the government did allow a network of private SRH care professionals to collaborate with international agencies and Non-Governmental Organizations (NGOs) to initiate a project aimed at filling this gap. Such a project could increase the prevalence of SRH facilities, offering affordable, accessible, quality services which, if successful, could increase awareness and knowledge of SRH, all with the desired outcome of behavioral change.

This project became the focus of the researchers. The project leader granted the students permission to interview key project members and stakeholders. The aim of the study was not to evaluate the offerings of the project as such, but to 'assess the potential indicators that determine success of the project as a case study.'

Prior to beginning their research, the master's students sought and received ethical approval from the local IRB, as well as from their own institutional ethical review board. Due to the sensitivity of the project, it was agreed that the

interviewees, the stakeholders, and the organization itself would remain anonymous, and all identifying information would be removed. All participants received an 'informed consent' agreement fully detailing the aims of the study. The agreement also contained a privacy statement that promised full confidentiality. All interviews were recorded, transcribed, and subsequently anonymized.

During the first few weeks of research, interviews were conducted on the participant's expectations, thoughts, and doubts surrounding the project. Many respondents demonstrated an acute awareness of the sensitivities regarding sexual and reproductive health. One stakeholder noted how 'increasing conservatism makes talking about SRH difficult,' and believed that professionals would be 'nervous raising these issues.'

In concluding their research, the master's students stressed the importance of the project for the community. They argued that although it touched upon sensitive issues, the project was neither illegal nor in violation of any state regulations. In order to make the project sustainable, it was recommended that 'partnerships between public and private sector need to be further developed,' and that perhaps 'business experts could be involved to establish a sustainable SRH service model.'

When a draft was presented to the SRH project leader, the students received word that there were still concerns about the 'potential harm' of their research. The students were told that they should consider removing all identifying information about the project from their report. While the project leader admitted that an ethical clearance had been issued by the local IRB earlier, and that promises of confidentiality had been met by the students, drawing attention to the sensitive issue of SRH services in a conservative culture with a taboo on sexual matters could have unforeseen and even adverse consequences for those involved, if not immediately, then perhaps in the future. Therefore, all names of the participants were either to be removed or anonymized, and any references to the actual project be omitted. Additionally, the report was to only be made public if it did not include a description of the SRH project.

This posed a dilemma for the students and their advisor. Firstly, it would be difficult to ensure the quality of their theses without describing the project being studied. Secondly, because their institution required that any master thesis project be submitted and subsequently archived at an institutional repository, it would therefore be made public and open for anyone to inspect in accordance with the scientific demand of transparency. Under the circumstances, it did not seem possible to fulfil participant requests for confidentiality *and* submit a Master's thesis in accordance with university requirements.

In practice, the requirement 'not to publish the report with the description of the project' would imply that the students could not formally finish their research project and thus not conclude their studies. The student's supervisor thereupon arranged for an exception to be made in this case, allowing the report to be archived without the possibility to inspect it, in turn effectively annulling the scientific merits of the study.

When we (the authors of this book) asked to report on this case, the students' supervisor allowed us access to relevant documents, including correspondence with

various stakeholders, on the provision that any and all identifying information be removed, including the names of the Master's students, the supervisor, and the SRH project leader, as well the name of the country where the research took place. After having completed our description, we destroyed all documents in our possession pertaining to this case were. Then we asked the supervisor and students involved to review this reconstruction, to see if they approved of it.

Assignment

Consider the nuances of this case. What efforts have the different parties (authors of this case study, project leader, supervisor, and students) pursued to ensure confidentiality? Do you believe these measures were enough? Can you think of another outcome that could have been possible had other steps been taken, and if so, what would you recommend? Discuss the case (and its possible alternative scenarios) in class, and answer the following questions:

1. Is it even possible to research 'sensitive issues' such as sexual and reproductive health, in a conservative culture without endangering the parties involved?
2. If so, what measures should be taken to ensure complete anonymity?
3. How, in the present situation, could a scientist still report on this project safely without giving up on academic freedom?

Suggested Reading

We recommend Sissela Bok's 1983 classic *The Limits of Confidentiality* and the chapter on confidentiality by Slowther and Kleinman (2008) for further orientation on the subject. We also recommend the chapter by Macnish (2020) in the *Handbook of Research Ethics and Scientific Integrity* (Iphofen, ed., 2020) for a discussion on the challenges inherent to privacy. Furthermore, we point to Manson and O'Neill (2007) for an extensive discussion on informed consent (though it is mainly focused on the medical sciences).

References

- Amiri, F., Yazdani, N., Shakery, A., & Chinaei, A. H. (2016). Hierarchical anonymization algorithms against background knowledge attack in data releasing. *Knowledge-Based Systems*, 101, 71–89. <https://doi.org/10.1016/j.knosys.2016.03.004>.
- Behnke, S. (2014, April). Disclosing confidential information. *Monitor on Psychology*, 45(4). <http://www.apa.org/monitor/2014/04/disclosing-information>.

- Bersoff, D. N. (2014). Protecting victims of violent patients while protecting confidentiality. *American Psychologist*, 69(5), 461–467. <https://doi.org/10.1037/a0037198>.
- Bjarnason, T., & Adalbjarnardottir, S. (2000). Anonymity and confidentiality in school surveys on alcohol, tobacco, and cannabis use. *Journal of Drug Issues*, 30(2), 335–343. <https://doi.org/10.1177/002204260003000206>.
- Blomley, N., & Davis, S. (1998). *Russel Ogden decision review*. Online: SFU President's Homepage, <http://www.sfu.ca/pres/OgdenReview.htm> (date accessed: 12 Mar 2020).
- Bok, S. (1983). The limits of confidentiality. *The Hastings Center Report*, 13(1), 24–31. <https://www.jstor.org/stable/3561549>.
- Buchmann, E., Böhm, K., Burghardt, T., et al. (2013). Re-identification of smart meter data. *Personal and Ubiquitous Computing*, 17, 653–662. <https://doi.org/10.1007/s00779-012-0513-6>.
- Ciriani, V., di Vimercati, S. D. C., Foresti, S., & Samarati, P. (2008). *K*-anonymous data mining: A survey. In C. C. Aggarwal & P. S. Yu (Eds.), *Privacy-preserving data mining. Advances in database systems*, vol. 34 (pp. 105–136). Berlin: Springer. https://doi.org/10.1007/978-0-387-70992-5_5.
- Dawson, P. (2014). Our anonymous participants are not always anonymous: Is this a problem? *British Journal of Educational Technology*, 45(3), 428–437. <https://doi.org/10.1111/bjet.12144>.
- Domingo-Ferrer, J., & Torra, V. (2005). Ordinal, continuous and heterogeneous *k*-anonymity through microaggregation. *Data Mining and Knowledge Discovery*, 11(2), 195–212. <https://doi.org/10.1007/s10618-005-0007-5>.
- Duncan, R. E., Hall, A. C., & Knowles, A. (2015). Ethical dilemmas of confidentiality with adolescent clients: Case studies from psychologists. *Ethics & Behavior*, 25(3), 197–221. <https://doi.org/10.1080/10508422.2014.923314>.
- El Emam, K., Jonker, E., Arbuckle, L., & Malin, B. (2011). A systematic review of re-identification attacks on health data. *PLoS One*, 6(12), e28071. <https://doi.org/10.1371/journal.pone.0028071>.
- Geraghty, R. (2016). *Anonymisation and social research*. Anonymising Research Data Workshop, University College Dublin, 22 June 2016. www.slideshare.net/ISSDA/anonymisation-and-social-research
- Giordano, J., O'Reilly, M., Taylor, H., & Dogra, N. (2007). Confidentiality and autonomy: The challenge(s) of offering research participants a choice of disclosing their identity. *Qualitative Health Research*, 17(2), 264–275. <https://doi.org/10.1177/1049732306297884>.
- Herbert, P. B. (2002). The duty to warn: A reconsideration and critique. *Journal of the American Academy of Psychiatry and the Law Online*, 30(3), 417–424. Retrieved from <https://pdfs.semanticscholar.org/5a4c/b550a640d165ec49c5a922291961c278aee6.pdf>.
- Hook, M. K., & Cleveland, J. L. (1999). To tell or not to tell: Breaching confidentiality with clients with HIV and AIDS. *Ethics & Behavior*, 9(4), 365–381. https://doi.org/10.1207/s15327019eb0904_6.
- Iphofen, R. (Ed.). (2020). *Handbook of research ethics and scientific integrity*. Cham: Springer.
- Irwin, S. (2013). Qualitative secondary data analysis: Ethics, epistemology and context. *Progress in Development Studies*, 13(4), 295–306. <https://doi.org/10.1177/1464993413490479>.
- Israel, M. (2014). *Research ethics and integrity for social scientists* (2nd ed.). London: Sage.
- Kaiser, K. (2009). Protecting respondent confidentiality in qualitative research. *Qualitative Health Research*, 19(11), 1632–1641. <https://doi.org/10.1177/1049732309350879>.
- Kelly, A. (2009). In defence of anonymity: Re-joining the criticism. *British Educational Research Journal*, 35(3), 431–445. <https://doi.org/10.1080/01411920802044438>.
- Khnlou, N., & Peter, E. (2005). Participatory action research: Considerations for ethical review. *Social Science & Medicine*, 60(10), 2333–2340. <https://doi.org/10.1016/j.socscimed.2004.10.004>.
- Kotch, J. B. (2000). Ethical issues in longitudinal child maltreatment research. *Journal of Interpersonal Violence*, 15(7), 696–709.
- Kumpošt, M., & Matyáš, V. (2009). User profiling and re-identification: Case of university-wide network analysis. In S. Fischer-Hübner, C. Lambrinoudakis, & G. R. Pernul (Eds.),

- Trust, privacy and security in digital business* (pp. 1–11). Berlin: Springer. https://doi.org/10.1007/978-3-642-03748-1_1.
- LaFrance, J., & Bull, C. C. (2009). Research ourselves back to life. Taking control on the research agenda in Indian country. In D. M. Mertens & P. E. Ginsberg (Eds.), *The handbook of social research ethics* (pp. 135–149). London: Sage. <https://doi.org/10.4135/9781483348971.n9>.
- Lelkes, Y., Krosnick, J. A., Marx, D. M., Judd, C. M., & Park, B. (2012). Complete anonymity compromises the accuracy of self-reports. *Journal of Experimental Social Psychology*, 48(6), 1291–1299. <https://doi.org/10.1016/j.jesp.2012.07.002>.
- Lowman, J., & Palys, T. (2000). Ethics and institutional conflict of interest: The research confidentiality controversy at Simon Fraser University. *Sociological Practice: A Journal of Clinical and Applied Sociology*, 2(4), 245–255. <https://doi.org/10.1023/A:1026589415488>.
- Lubarsky, B. (2017). Re-identification of “anonymized data”. *Georgetown Law Technology Review*, 202, 202–213. <https://perma.cc/86RR-JUFT>.
- Macnish, K. (2020). Privacy in research ethics. In R. Iphofen (Ed.), *Handbook of research ethics and scientific integrity* (pp. 233–249). Cham: Springer. <https://doi.org/10.1007/978-3-319-76040-7>.
- Manson, N. C., & O'Neill, O. (2007). *Rethinking informed consent*. Cambridge: Cambridge University Press.
- Martin, D. J., Kifer, D., Machanavajjhala, A., Gehrke, J., & Halpern, J. Y. (2007, April). Worst-case background knowledge for privacy-preserving data publishing. In *2007 IEEE 23rd international conference on data engineering* (pp. 126–135). Piscataway: IEEE. <https://doi.org/10.1109/ICDE.2007.367858>.
- Munson, R. (2008). *Intervention and reflection: Basic issues in medical ethics* (8th ed.). Belmont, CA: Thomson Wadsworth.
- Narayanan, A., & Shmatikow, V. (2009). De-anonymizing social networks. In *2009 30th IEEE symposium on security and privacy* (pp. 173–187). Los Alamitos: IEEE. <https://doi.org/10.1109/SP.2009.22>.
- Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review*, 57(6), 1701–1778. <https://ssrn.com/abstract=1450006>.
- Patel, D. (2016). Research data management: A conceptual framework. *Library Review*, 65(4/5), 226–241. <https://doi.org/10.1108/LR-01-2016-0001>.
- Ramachandran, A., Singh, L., Porter, E., & Nagle, F. (2012). Exploring re-identification risks in public domains. In *2012 tenth annual international conference on privacy, security and trust* (pp. 35–42). Paris: IEEE. <https://doi.org/10.1109/PST.2012.6297917>.
- Rhoen, M. H. C. (2019). *Big data, big risks, big power shifts: Evaluating the general data protection regulation as an instrument of risk control and power redistribution in the context of big data (doss.)*. Leiden: Leiden University. <https://openaccess.leidenuniv.nl/handle/1887/77748>.
- Samarati, P., & Sweeney, L. (1998). *Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression*. Technical report, SRI International. Retrieved from: https://epic.org/privacy/reidentification/Samarati_Sweeney_paper.pdf
- Scarce, R. (1995). Scholarly ethics and courtroom antics: Where researchers stand in the eyes of the law. *The American Sociologist*, 26(1), 87–112. <https://doi.org/10.1007/BF02692012>.
- Scott, R. C. (1995). Anonymity in applied communication research: Tension between IRBs, researchers, and human subjects. *Journal of Applied Communications*, 333, 242–257. <https://doi.org/10.1080/00909880500149445>.
- Slowther, A., & Kleinman, I. (2008). Confidentiality. In P. A. Singer & A. M. Viens (Eds.), *The Cambridge textbook of bioethics* (pp. 43–50). Cambridge: Cambridge University Press.
- Taddei, S., & Contena, B. (2013). Privacy, trust and control: Which relationships with online self-disclosure? *Computers in Human Behavior*, 29(3), 821–826. <https://doi.org/10.1016/j.chb.2012.11.022>.
- Thomson, D., Bzdel, L., Golden-Biddle, K., Reay, T., & Estabrooks, C. A. (2005). Central questions of anonymization: A case study of secondary use of qualitative data. *Forum: Qualitative Social Research*, 6(1), Art. 29. <http://nbn-resolving.de/urn:nbn:de:0114-fqs0501297>.

- Tilley, L., & Woodthorpe, K. (2011). Is it the end for anonymity as we know it? A critical examination of the ethical principle of anonymity in the context of 21st century demands on the qualitative researcher. *Qualitative Research*, 11(2), 197–212. <https://doi.org/10.1177/2F1468794110394073>.
- Tolich, M. (2004). Internal confidentiality: When confidentiality assurances fail relational informants. *Qualitative Sociology*, 27(1), 101–106. <https://doi.org/10.1023/B:QUAS.0000015546.20441.4a>.
- Vainio, A. (2013). Beyond research ethics: Anonymity as ‘ontology’, ‘analysis’ and ‘independence’. *Qualitative Research*, 13(6), 685–698. <https://doi.org/10.1177/2F1468794112459669>.
- Van den Hoonaard, W. C. (2003). Is anonymity an artifact in ethnographic research? *Journal of Academic Ethics*, 1(2), 141–151. <https://doi.org/10.1023/B:JAET.0000006919.58804.4c>.
- Walford, G. (2005). Research ethical guidelines and anonymity. *International Journal of Research & Method in Education*, 28(1), 83–93. <https://doi.org/10.1080/01406720500036786>.
- Weinberg, M. (2002). Biting the hand that feeds you and other feminist dilemmas in fieldwork. In W. C. van den Hoonaard (Ed.), *Walking the tightrope: Ethical issues for qualitative researchers* (pp. 79–94). Toronto: University of Toronto Press.
- Whelan, T. J. (2007, October). *Anonymity and confidentiality: Do survey respondents know the difference?* Poster presented at the 30th annual meeting of the Society of Southeastern Social Psychologists. Durham, NC.
- Wiles, R., Charles, V., Crow, G., & Heath, S. (2006). Researching researchers: Lessons for research ethics. *Qualitative Research*, 6(3), 283–299. <https://doi.org/10.1177/2F1468794106065004>.
- Williams, G., & Pigeot, I. (2017). Consent and confidentiality in the light of recent demands for data sharing. *Biometrical Journal*, 59(2), 240–250. <https://doi.org/10.1002/bimj.201500044>.
- Zang, H., & Bolot, J. (2014). Anonymization of location data does not work: A large-scale measurement study. *2014 IEEE International Conference on pervasive computing and communication workshops (PERCOM WORKSHOPS)*, Budapest, Hungary, 24–28 March 2014. <https://doi.org/10.1145/2030613.2030630>.
- Zhou, B., Pei, J., & Luk, W. S. (2008). A brief survey on anonymization techniques for privacy preserving publishing of social network data. *ACM Sigkdd Explorations Newsletter*, 10(2), 12–22. <https://doi.org/10.1145/1540276.1540279>.
- Zimmer, M. (2010). ‘But the data is already public’: On the ethics of research in Facebook. *Ethics and Information Technology*, 12(4), 313–325. <https://doi.org/10.1007/s10676-010-9227-5>.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.