

14

THE RESEARCH PROCESS

Key questions

- What are the key steps of the research process?
- How can the researcher move from a general idea about what to study towards formulating a workable question to guide the research?
- Which is the best strategy for deciding on a unit or field of study – or several units and fields – and how might they be effectively delimited?
- What are some of the key challenges faced by the researcher in gathering data to form the basis for empirical analysis?
- What are the key elements of data analysis?

Key concepts

Emergent research questions * multi-sited research * researching sites, topics, people * research logging * saturation * thematic analysis

With this quite hands-on methodological chapter, I present a framework for digital social research. The chapter provides a step-by-step discussion of the research process, and focuses on formulating a good research question, choosing your data, collecting it, analysing and interpreting it, as well as arriving at research results. The research process, as described here, relies on tools from digital ethnography (detailed in Chapter 15) and also on a set of approaches to mapping and mining

digital sociality (detailed in Chapter 16). As I have already discussed in Chapter 13, and as will be further discussed in Chapter 15, I advocate an approach that rests on the solid foundation of ethnography. This means taking a holistic and interpretative approach to whatever aspect, dimension, or characteristic of digital society that you want to study. Drawing on different forms and levels of observing, taking part in, experimenting with, discovering, and mapping your object of analysis is a vital part of the strategy.

Depending on what you want to analyse, what questions you aim to answer, and which claims you want to be able to make, your ethnographic approach may vary in its character. This may be to the point where some researchers might claim that it is in fact no longer deserving of being called ‘ethnography’ in the sense that it draws on conventional field data and interpretative techniques. It may range from undertaking traditional forms of anthropological fieldwork for extended amounts of time, relying solely on ethnographic data and ‘qualitative’ methods, to just being ethnographically inspired by the aim to make sense of people’s interaction in social contexts, but relying just as much on types of data and analysis that are not seen to be conventionally ethnographic. This is in line with marketing researcher Robert Kozinets’ concept of what he calls *netnography*, which he says is something quite different from standard ethnography. However, this does not mean that any strategy whatsoever can count as research. Kozinets (2015: 101) writes:

Are you ready to begin a netnographic research project? Pull that smartphone out of your purse or pocket. Flip open your laptop or uncover your handy dandy tablet. Double tap that app. Type some keywords into that search engine, pick your site and you should be good to go. You are just a few clicks away from finding a fabulously free-flowing online conversation about just about anything. And then you are off, entering the wonderful world of netnography. Or maybe not.

Rather, before one can start researching — interacting, discovering, recording, and reflecting — it is important to decide not just what one is going to study, but also how. We must be able to respond to crucial questions about the who, what, where, and how of our empirical work. We will look closer at Kozinets’ approach in the next chapter, but what is important for the time being is his notion that one must have a general sense of the response to such questions, rather than having things set in stone. Discussing ethnography in internet research, danah boyd argues that one’s methods should be a customised, evolving set of approaches and strategies. The main goal is to get a holistic picture of the object of analysis. She writes:

While approaches vary, ethnographic fieldwork relies on participant observation, qualitative interviews, and analysis of cultural artifacts to make sense of cultural

practices on their own terms. Ethnographers use different techniques to interpret, complicate, and analyze cultural practices, situate complex cultural phenomena, and map social worlds from the bottom up. As a method, ethnography does not speak to individual traits or beliefs but to the complexity and interconnectedness of culturally driven practices and norms. Ethnography produces a topological map of a particular set of cultural practices. (boyd 2008: 46–47)

So the thing that the researcher should aim for is that *topological map*. And to capture it, one must rely on a set of different techniques to analyse and situate the data. This is why the concept of ‘methodological bricolage’, discussed in Chapter 13, is so useful. So, once again, even though I believe that the ethnographic approach is very well suited to unravel how digital society — or part of it — functions, the emerging data environment (see Chapter 13), as well as the complexity of social life, demands that one considers multiple techniques to broaden the investigation. Such techniques, as will be further presented in Chapter 16, may be, but are not limited to, ‘following the medium’, social network analysis, and text mining.

While observation and interviews are central to ethnographic research, it is common that ethnographers augment those methods with other types of data collection. It is all about being pragmatic and adaptive. Anthropologist Tom Boellstorff (2010: 129) also agrees that the research method for any project is constantly renegotiated and transformed. He writes:

Any claim that a particular method is the best (or the only valid) method [...] misses how research always involves a coming-together of research question and methodology. How one conducts research is not determined by some essential property ‘out there;’ it is determined by the research questions that one wants to investigate.

FORMULATING THE RESEARCH QUESTION

The research process, in most kinds of inquiry, tends to be iterative and reflexive, which leads to revisions of the choices that have been made, sometimes several times. But thinking about *the research question* is a good place to start. What is it that you want to study? Deciding on a research question, or a set of questions, is a fundamental and consequential part of the research process. This is because it is the decision from which most other challenges and choices throughout the research will follow. Once the research question is in place, at least an initial and preliminary one, data collection can be started, guided by that question.

In order to arrive at a research question, it is a good idea to start with some more general questions and drill down from there. Broad questions about the general

area in which we take an interest will help us to explore the setting or topic we are interested in. Discussing her ethnographic work on youth and social media, boyd (2008: 44) reflects upon how her research question was repeatedly revised during the research process:

My ongoing interest in identity, privacy, impression management, and social interaction shaped where I began, but my fieldwork also led me to explore other topics. While I had grown up at a particular point in the history of the Internet, it was tremendously rewarding to be able to watch a new generation of teenagers embrace an entirely different set of circumstances and technologies. I was well positioned to watch the phenomenon of teen engagement with social network sites unfold and I used that to my advantage. MySpace became popular with teens just as I began my fieldwork. I did not initially intend to focus on social network sites, but as teens turned their attention to these sites, so did I.

So, research questions are emergent. It is common that one starts off a project with a set of questions that are subsequently revised so that one comes out the other side with a quite different set of questions from the ones that were there at the beginning. While part of this process is about going along with the field, as boyd did, another part is about looking carefully at any previous research that has been done in this or related areas. Doing a literature review allows you to get a good picture about the domain to which you want to make a contribution, and mapping the state of the field will help sharpen your research interests. Boellstorff and colleagues (2012: 55) encourage researchers to ‘just plunge in and start reading’. Remember to think historically too about studies in related — even pre-digital — fields that may be relevant to frame your own work, and to think widely and to explore studies from less obvious parts of the research landscape. When it comes to the need to narrow down your research question, Kozinets (2015: 122) has a set of practical suggestions (drawing on Creswell 2009: 129–131):

- 1 Formulate a single large, broad, guiding question (you can always narrow it down later).
- 2 Ensure that your question is amenable to netnographic inquiry. That is, does your question relate to the online social interaction and experience of particular sites, topics or people? If not, go back to the drawing board to reformulate.
- 3 Building on that large question, formulate no more than five related sub-questions that elaborate parts of your major question.
- 4 Try to focus on the question word you are using. If you are interested in people and topics or their locations in online sites, then ‘where’ may be important. If you are interested in processes, then ‘how’ and ‘when’ may be important. If you are

interested in people who inhabit sites or discuss topics, then ‘who’ may be important. Your most useful questions often will begin with ‘what’, as they relate to descriptions of things such as types of online narrative, story, topic, meanings, or associations. ‘Why’ tends to be a very difficult question for netnographic evidence to conclusively answer.

- 5 Try experimenting in your question with exploratory verbs such as ‘discover’, ‘understand’, ‘explore’, ‘describe’, or ‘report’.

The general idea is to arrive at the research question through *exploration*. Look around, test ideas, discard some, and hold on to others. However, this openness should not be mistaken for ‘wandering around aimlessly’, as having an adaptive approach still means that your work must be purposive. Many of the phenomena that we study in digital society are relatively novel, emergent, and shift in their character, which demands that the researcher is prepared to modify questions based on what is encountered throughout the study. Back in the 1920s, anthropologist Bronislaw Malinowski (1922: 9) argued that the work of the ethnographer (just like that of any researcher) is ‘worthless’ unless he or she is prepared to be ungrudgingly casting any determination to prove certain hypotheses aside ‘under the pressure of evidence’. So, the research question is often formulated *iteratively* throughout ‘an emergent process of discovery’ (Boellstorff et al. 2012: 54).

Research questions can be formulated on different levels of abstraction, and they can be more or less specific. Drawing on the previous chapters in this book, one can imagine a wide range of questions that would be interesting, relevant, and important to study, as parts of the wider inquiry into digital society. One such question could be: In what different ways are emotions and social support expressed in open and anonymous forums as compared to in closed and non-anonymous ones? Another one could be: What parts of their everyday lives are people most likely to share with others in the form of photos posted on social media? Yet another could ask: To what extent and how is hate speech countered when it occurs online? Or: How do ‘amateur’ creators of online content feel about posting their creations on privately owned profitable platforms? No matter what our question happens to be, and what specificities about digital sociality, culture, and interaction it is interested in, the mark of a really good research question is that it centres on issues that are important in relation to wider research communities (Boellstorff et al. 2012). In general, it is a very good idea to make your research narrow while keeping your thoughts broad, which means formulating the question in a way so that it has the potential to produce results and generate knowledge that is relevant to larger debates in society. For example, we might study some aspects of online gaming and — nearly as a bonus — be able to say something more generally about, for example, intimacy and friendship in the 21st century. We may analyse social network patterns in a feminist discussion forum online, and also get to know new and important things about social movements and gender in today’s world. A really

good question will, if the research is successful, generate knowledge that is useful in more than one way. Boellstorff and colleagues (2012: 56) make a very important point when they say that it can't be overemphasised that, aside from its relevance, a research question is only as good as the researcher is passionate about it:

All good science flows from a scientist's passion to learn something he or she is deeply curious about. Thus [...], although it is important for the work to be broadly relevant, the question should be personally interesting and exciting to us. It is vital to underscore this point because we are best served by honoring our own passions and intellectual journey when deciding on a research question.

FRAMING THE FIELD

Once the initial research question is set, the next step is to decide what constitutes the core unit of study. Which setting, which group, or individuals will you study, and what are your delimitations in space and time. Basically: How should you frame your 'fieldsite'? The concept of fieldsite will be discussed closer in Chapter 15, but for now it suffices to say that it is a name for the context you want to study. Boellstorff (2010) has been working with ethnography in the specific settings of virtual worlds — such as online multiplayer games and other online social realms that aim to render some sort of 'world' for users to be embedded in — but has developed thoughts and guidelines that are just as relevant for any form of ethnography in digital society. Part of Boellstorff's thinking is a three-part typology of what he calls ethnographic scale:

- First, he writes, we might be interested in focusing on one specific and clearly delimited online setting as a fieldsite in its own right.
- Second, we might want our study to explore — and maybe compare — more than one such relatively distinct setting.
- Third, we may want to move in the direction proposed by Daniel Miller and others (see Chapter 15) and analyse not only online settings but focus — maybe even wholly — on the role of one or several online settings in the 'actual' lives of people. This means starting the study offline.

Kozinets (2015) makes a similar distinction between frames for studies. He writes of a scenario synonymous to Boellstorff's first one:

We might be interested in examining a particular online community, for example, Wikipedia, [the newsgroup] alt.coffee, the Winter is Coming Game of Thrones news blog, the use of #addiction hashtags, or the Church of Satan's official Facebook group. (Kozinets 2015: 118)

He then goes on to write about the second, and *multi-sited*, and also the third 'offline' scenario:

Alternatively, a research topic can be considered to be manifesting in a widely dispersed manner among a wide number of online and offline social experiences. A site such as a Star Trek or Samsung Galaxy wiki can seem at first glance to be a straightforward single location. [...] However, [...] my interest in a particular topic, such as Star Trek fandom, drew me to a variety of sites or groups. Sometimes, certainly, my entire study stayed on one site such as alt.coffee, but this was as much topical as it was a response to the relative concentration of the early Internet prior to the explosion of blogs and other social media forms. (2015: 118)

Similarly, boyd (2008: 54) writes that her research on teens and social network sites also saw her develop a 'multi-sited project' where her fieldsite, in practice, was a network of several different sites. In her case, these sites were both online and offline, which is quite often the case. It is common that the analysis of social things on the internet will urge the researcher to look for information in non-internet places as well in order to get a more complete picture. Furthermore, while Boellstorff's three-tiered terminology can indeed be helpful to think about how we delimit and define our fieldsite, the boundaries between sites both online and offline, and between online and offline as such, are perpetually unstable. Because of this, many studies end up encompassing both of these dimensions. As in boyd's case, the researcher may move between the online and the offline simply by following the field where it leads. At other times, researchers may design their studies from the outset to pay explicit attention to offline events and concerns.

It may also be possible to frame the context to be studied, but on the basis of investigating an *activity*, rather than in terms of milieu and setting, or a particular group and the way it functions. It may be helpful to think in terms of the idea, presented by Kozinets (2015), that ethnographic research is based on studies of *sites*, *topics*, and *people*:

- Sites, then, are locations — any kind of social site, geographically, culturally, or notionally. These are the places where people's social worlds take root and it is the job of the ethnographer to map them out. Research into such sites respond to questions of *where*.
- Topics are conceptual, rather than spatial, and are related to language, knowledge, information, and meaning. Topics are sites of attention and researching them helps the researcher respond to questions of *what*.
- People are what we analyse when we want answers to questions about *whom*.

It can be useful to think of one's research in relation to concepts and divisions such as these. Not only is it necessary to be able to break down and describe the research process to others afterwards in a comprehensible way, but it is also good to gain perspective on one's own relationship to the object of study, because in practice the levels and dimensions float into one another. Reflecting on this, boyd (2008: 55) describes her process:

Instead of starting with one bounded site, I decided to approach my field site as a network. I focused my study on the intersection of American teenagers, their relationship to networked publics, and, in particular, the sociotechnical phenomenon marked by the rise of social network sites. I began my fieldwork from different angles and traversed the phenomenon using different approaches. My fieldwork includes mediated and unmediated environments and I moved across different social contexts and engaged with different relevant social groups to gain an understanding of what was taking place. Approaching this puzzle, I began broadly and narrowed my focus as I achieved clarity. As appropriate, I expanded my scope when following specific people or trying to make sense of specific spaces. This created many layers of awareness that allowed me to locate people, spaces, and practices in a broader context.

COLLECTING DATA

After thinking about the field of the study, the next step is to collect the data to be analysed. As outlined in Chapter 13, these data can be of varying character. We are working on an ethnographic foundation here, so therefore 'qualitative' data (observation notes, interview transcripts, and the like) will play an important role. As I have argued earlier, such data is enough in some cases. Sometimes (see Chapter 16), it will be useful to complement them with data which emanates from natively digital strategies for collection and analysis, what Rogers (2013) calls 'methods of the medium' such as search engine results, likes, links, tags, and so on. We may also be interested in 'quantitative' data to use in social network analysis, or large sets of downloaded text data to analyse with the help of 'quantitative' text mining techniques (more about this in Chapter 16, as well). So, as discussed in Chapter 13, the emerging and increasingly complex data environment means that any researcher approaching digital society — or slices or bits of it — will want to access different data types that demand different techniques to capture them. Observational data can be registered through note taking, interviews may be recorded and transcribed, or done through email or text messaging, and will be automatically captured as text in that way. Digitally native data, social network data, and big text data-sets will have to be captured through any means available — copying and pasting when possible, downloading when possible, and by scraping (using tools to automatically

download it) or acquiring data from a variety of online services, databases, websites, and resources according to different schemes from case to case. Part of becoming an empirical researcher of digital society is to embark on the learning journey of discovering tools, strategies, and resources — mastering them, and constantly adapting to their variable and changing functionality, availability, and terms of use.

Similarly to the steps involved in the formulation of the research question, and framing the field, the step of data collection also lends itself to an open and explorative starting phase. Kozinets (2015: 167–168) describes this process in an engaging way:

I am often asked which special software tool to use. Try everything out there, online and offline. [...] For the sake of illustration, you can simply use Google — and this includes additional Google features such as Google's Analytics, Trends, and NGram reader features. [...] Use several engines from Google, including the web, groups, blog and image search function. Then search YouTube videos. Search on Twitter, and on a forum search engine like omgili.com ('OMG I love it!'). Take a look at Facebook Groups, Wikis and LinkedIn groups. [...] Keep your search terms as simple and consistent as possible across sites and engines to start. Grab as many major overviews of data as you can, but always also stay as close to the interactions and direct experiences as possible. [...] Your main task at this stage is to first attune your perception to the various social media and other channels that might inform your research question. [...] For example, if you are studying the contemporary whale and seal hunts worldwide, then consider entering variations on 'whaling', 'whale hunt', 'whale activism', 'whale management', 'conservation', 'hunting endangered species', 'animal rights', 'Native rights', 'Aboriginal rights', 'international markets for whale meat and blubber' and 'international hunting accords dating back over 1000 years'. Investigate everything. Look at every website that seems even remotely relevant. Read them, and follow all the trails and hints they bring to you. Take your time. [...] In the long run, you must choose particular routes and pursue those, seeing where they take you. [...] How and why you direct your research will be driven both by your research questions as well as by serendipitous discoveries that you make along the way.

After the exploratory phase, which may or may not be as wide-reaching and ambitious as that which Kozinets describes, a focused and structured phase of data collection — guided by the research question, and targeted towards the chosen object of study (sites/topics/people) — must commence. As mentioned above, the actual process of doing this will probably include several different strategies, depending on the type of data. One important thing is to document your research thoroughly: 'If we fail to write it down, it might as well not have happened!' (Boellstorff et al. 2012: 82). So, it is important to

take notes throughout the process, as well as keeping data organised. Make it a habit to keep a *research log*, where you note things such as the search queries you have made, what settings you manipulated and how, what day you made a research interview, which filter and layout settings you used in your social network analysis software — all of these things are useful to note. You will thank yourself for having kept that log when the time comes to write about your method when you are about to publish the results of your research. From the very start of the data collection process, observational data, interview data, and/or data gathered through any other methods must be neatly logged. Some researchers do this in notebooks, others in word processing software, and others use databases. You need to take this type of note taking and documentation seriously. Boellstorff and colleagues (2012: 113) write that the handy possibilities for copying, pasting, and downloading data in online contexts creates a risk that other, researcher-centred, forms of documentation are forgotten:

This ease of data capture can sometimes create the false impression that the methods enumerated earlier, such as taking fieldnotes based on participant observation, are unnecessary. Why go to all that trouble when we can just capture the data digitally? This temptation is false: we cannot stress enough that as useful as digital records are, they cannot stand alone without the rigor of a detailed accounting of interactions in the field.

FURTHER READING ABOUT CAPTURING AND COLLECTING DATA

Emerson, Robert M., Fretz, Rachel I., & Shaw, Linda L. (2011). *Writing Ethnographic Fieldnotes*. Chicago, IL: University of Chicago Press.

James, Nalita, & Busher, Hugh (2009). *Online Interviewing*. London: Sage.

Russell, Matthew A. (2014). *Mining the Social Web*. Sebastopol: O'Reilly Media.

Salmons, Janet (2016). *Doing Qualitative Research Online*. London: Sage.

Data collection in digital social research can and should be open, exploratory, and aim to get as complete an image as possible. It is useful to think about the concept of *saturation* when it comes to the issue of when one should stop data collection and move on with analysis. This concept, originally formulated by the originators of the so-called

Grounded Theory method, Barney Glaser and Anselm Strauss (1967), postulates that once the research process has been going on for a while, the researcher will gradually develop a sense of whether a newly added piece of data points to any new insights. If the answer is yes, then the data collection continues. If the answer is no, the data is disregarded, since it only adds bulk to the dataset and nothing to the theoretical insights (Glaser 1965: 441–442). So, as Boellstorff and colleagues (2012: 59) write:

When we start hearing the same reflections repeated in interviews, when we are no longer seeing new things or getting new insight while undertaking participant observation, when we have reached a point where we can anticipate answers, practices, and the general everyday unfolding of the field, we have likely reached the point of diminishing returns in our data collection and can consider that phase complete.

Basically, when we feel that we have enough data to be able to say something that is meaningful and interesting, and when our sample is sizeable enough to provide a solid foundation for the arguments that we want to make, then it is time to turn to the analysis phase.

ANALYSING AND INTERPRETING DATA

In the next step, the data that we have collected is used as the raw material for analysis and interpretation. Analysis is what we do when we break the studied phenomenon down into its component parts. Boellstorff and colleagues (2012: 166) explain that ‘the key to data analysis is to interact with the dataset: read it, study it, immerse oneself within it, and let the data paint a portrait of the culture we are studying’.

As you will be able to tell from the approaches to be presented in Chapters 15 and 16, the methods for analysing data can indeed differ. On the one hand, they can be resting on a statistical foundation, as is the case in, for example, social network analysis and text mining. On the other hand, they can be interpretative, as in the cases when it draws on ‘qualitative’ observational or interview data. In line with this, Kozinets (2015: 198) argues that we should be open to using both hermeneutic interpretation, and all sorts of computational elements. While it will be further detailed in Chapter 16 what social network analysis and text mining can entail, there is reason here to think more about what hermeneutic interpretation is. Kozinets (2015: 205) gives an example:

Consider a collection of 3 YouTube videos, 4 pages of blogger text, 6 pages of newsgroup materials and 17 Instagram posts. Finding their common elements, the key and core of their meaning structure, requires us to find the common elements between them. Locating these shared themes is the challenge of hermeneutic interpretation.

Hermeneutics is an interpretative strategy where parts and instances of the data are iteratively reassessed and interpreted in relation to a developing sense of what the data say as a whole. So, one starts to look at some of the data to arrive at some initial understandings. Those understandings inform subsequent readings of other parts of the data, and this process carries on in interpretative loops, while the researcher's understanding of the meaning of the entire context that is represented by the dataset takes shape. But, while that rather abstract and philosophical notion of how knowledge is formed through empirical experience is a useful and inspirational image, we also need actual hands-on strategies to move the analysis forward to broader insights.

First, the data that has been collected must be systematised through some form of *thematic analysis* (see for example, Braun & Clarke 2006). The practice of annotating data, assigning tags or keywords to important passages and insights, is called *coding* and it is something that is done early on in the research process to map prominent themes in the empirical material. Moving on from reading the data, annotating it, promoting some annotations into codes, one is looking to develop more abstracted themes. In sum, the analysis should move from unstructured data, to annotated and coded data, to themes.

FURTHER READING ABOUT CODING AND THEMATIC ANALYSIS

Guest, Greg, MacQueen, Kathleen M., & Namey, Emily E. (2012). *Applied Thematic Analysis*. Thousand Oaks, CA: Sage.

Saldaña, Johnny (2015). *The Coding Manual for Qualitative Researchers*. Thousand Oaks, CA: Sage.

Once one has arrived at the core themes of the context under analysis, the next and final step of the study is to develop arguments and conclusions that point to larger theoretical and conceptual insights. By necessity, the scope and type of the results that will emerge from your research will range from the narrow to the overarching, from the abstract to the concrete, from descriptions to applicable insights, all depending on what the study is about, and on the methodological choices made along the way. 'Ethno-graphy' means culture writing, and the primary output of such research is the documentation and hermeneutic interpretation of the social setting that has been analysed. So, at the end of the research process, when insights have been channelled

through codes and themes into broader insights that can be developed, discussed, and positioned in relation to previous research and to theoretical concepts and debates, it is time to start crafting your text. In fact, the writing process is also part of the inquiry as one often arrives at important analytical insights through the process of writing in itself.

FURTHER READING ABOUT RESEARCH WRITING

Sword, Helen (2012). *Stylish Academic Writing*. Cambridge, MA: Harvard University Press.

Wolcott, Harry F. (2009). *Writing up Qualitative Research*. Los Angeles, CA: Sage.

Digital Social Research — some pointers

- 1 During any social research, especially that which explores 'new' or uncharted phenomena, always strive to capture as much complexity as possible. Ethnography — striving for 'thick description' (see Chapter 15) is a good basic philosophy, even though it can be scaled and adapted largely depending on the character of your study.
- 2 Drawing on the ethnographical basis, bring in other sources and techniques for getting a topological map of the analysed phenomenon that is as rich as possible. In studying the internet and social media, you can consider methods like social network analysis and text mining (see Chapter 16).
- 3 After initial exploration, formulate a research question that is as concrete and sharp as possible. Don't worry if you have to revise it continuously throughout doing the research. Aim to formulate the question in such a way that the results of your study will potentially be useful beyond your specific empirical case.
- 4 Decide what you are going to study. Is it sites, topics, or people, or more than one of these? Will your study be focusing on one delimited setting, or will it be multi-sited? What role will the 'offline' play in your study?

(Continued)