

Exploring the Meaning Problem of Big and Small Data Through Digital Method Triangulation

Sara Leckner^I & Pernilla Severson^{II}

^I Department of Computer Science and Media Technology, Faculty of Technology and Society, Malmö University, Sweden, sara.leckner@mau.se

^{II} Department of Media and Journalism, Faculty of Arts and Humanities, Linnaeus University, Sweden, pernilla.severson@lnu.se

Abstract

In this article, knowledge building through combinations of methods in a digital context is discussed and explored. Two types of digital bigger and smaller data-driven media studies are used as examples: digital focus groups and the combination of internet traffic measurements, surveys and diaries. The article proposes the concept of digital method triangulation. Digital method triangulation is argued to be a way to approach the “meaning problem” to make sense of small and big data. Digital method triangulation is argued 1) to stimulate the innovative use of known methods for unexpected dimensions within the studied topic; 2) with appropriate theoretical and meta-theoretical reflections, to provide more certainty in conclusions; and 3) to assist in constructing a more comprehensive perspective on specific analyses. The conclusion is that triangulation is even more important in the digital realm, as it facilitates dialogue between conventional and digital methods, dialogue that seems crucial to capture the complexities of the onlife.

Keywords: digital methods, triangulation, meaning problem, methodological development, big and small data

Introduction

The meaning problem – sometimes understood as the inference problem or frame problem – of big and small data is an important issue in critical data studies (Iliadis & Russo, 2016; Kitchin, 2014; Kitchin & Lauriault, 2014; Lupton, 2014; Manovich, 2012). Deducing from data, understanding the world from data rather than from theory, is seen as problematic, because data come from somewhere, are always partial and are part of particular contexts and interests (Kitchin, 2014). The meaning problem is related to a critically engaged method discussion on research practices regarding what big data are

Leckner, S. & Severson, P. (2019). Exploring the meaning problem of big and small data through digital method triangulation. *Nordicom Review*, 40 (Special Issue 1): 79-94. doi:10.2478/nor-2019-0015.

and how to find meaning in digital media data (Mahrt & Sharkow, 2013). It is often expressed as the problem of interpreting data, questioning the meaning of observed data and analysing data. This includes an awareness of the online–offline contexts as really an onlife setting, that is, as being “inextricably interwoven” with each other (Simon & Ess, 2015: 157). The ontological distinction between the natively digital and the digitized has led to the question of whether it should also apply to the methods used to understand the meaning problem. Following Rogers’s (2013) paraphrasing of virtual and digital methods, is the digital best studied through *digital* methods, that is, methods that use the epistemology of the internet as a methodological basis, or are *virtual* methods, that is, digitalized conventional methods, still necessary to capture the understanding of the digital? A related issue is the implicit assumption that the meaning problem of (big) data studies only implies large quantities of data. There is an increasing discussion about how small data offer an important alternative to big data and provide meaning to big data, as they can be perceived as more insightful, detailed and manageable in size (Kitchin & Lauriault, 2015; Lupton, 2014). Furthermore, and related to the prior challenges, there is the meaning problem in relation to observed behavioural data and self-reported data. In the digital and big data setting, data traces can be understood as non-obtrusive and truer because they are a by-product of the everyday behaviour of users, which can ensure a certain degree of ecological validity (Mehl & Gill, 2010). However, it also involves asking meaningful questions and being aware of how the choices made, at any given point in the research process, will affect all the subsequent phases. This requires a strong need for theoretical reflection, in clear contrast to the alleged “end of theory” (Anderson, 2008; Bailenson, 2012; Mahrt & Sharkow, 2013). Thus, seeing that the meaning making of digital traces in the onlife is a complex methodological process, which has been argued to be the most effective when the researcher takes account of it (boyd & Crawford, 2012), such an approach is proposed in this article, based on what has been argued to be the key to conducting critical big data studies, namely combinations of methods (Kitchin, 2014; Lupton, 2014).

The present article argues and describes how data-driven media studies explore combinations of methods in the digital domain to make sense of human traces in the onlife and, by using two such examples, can be strengthened by increased methodological awareness, involving a critical (big) data perspective. For this purpose, the article proposes the concept of digital method triangulation as a collective term to incorporate digital structuring, conditions and meanings of media studies of the onlife and demonstrates that the concept can be a valuable approach to methodological development in the field of social science and humanities.

Digital method triangulation is proposed to be able to include combinations of qualitative and quantitative approaches of offline and online methods as well as of big and small data. Therefore, the two types of data-driven media studies used as examples – 1) digital focus groups and 2) measurements of internet traffic data, surveys and diaries – include such combinations. The two examples are well suited to making sense of big and small data. They address different aspects, problems and challenges that occur in the studying and sense making of the onlife yet are still not excessively utilized. The examples particularly show the interrelationship between offline and online grounded methods as well as their different ontological and epistemological bases. The onlife approach of meaning making explores exactly this (e.g., Simon & Ess, 2015).

Although method triangulation is not a new phenomenon, digitalization has aroused a need for the combination of methods to capture the breadth of practitioners and values created by new technologies. However, since the emerging, growing and maturing field of making sense of large and small data is highly interdisciplinary, there are many possibilities for new methodological synergies to capture this complexity; at the same time, it has resulted in debates and power struggles, requiring critical interrogation of assumptions and prejudices (see e.g., boyd & Crawford, 2012). The aim of this article is thus to contribute to an increased awareness of how media scholars can develop their knowledge of method triangulation to obtain meaningful inferences and interpretation of big and small data in the onlife context.

The meaning problem of big and small data

The meaning problem of big and small data is not self-evident. A highly interdisciplinary field such as critical data studies can benefit from clarification of the key concepts. In the broad field of social sciences and humanities, the digital and the meaning problem have been understood, discussed and debated in various ways.

Within established disciplines, like sociology and anthropology, sub-research fields focusing on the digital have evolved. They have approached the meaning problem as part of the interplay of embeddedness and negotiations with digital media in everyday life, including analyses of digital use as well as digital data analysis. The sub-research field of digital sociology has used sociological concepts and approaches to understand the digital in society (see e.g., Lupton, 2014). Marres (2017), for example, related digital sociology to other fields, like media and communication studies and computer science. The meaning problem has been discussed in this setting as an epistemological and therefore methodological question. Methods are actually “interface methods”, parts of a digital infrastructure as well as able to be configured and reconfigured (Marres, 2017). In the sub-discipline of digital humanities, the meaning problem is understood as the inability for research to move from data to arguments and interpretations (Arthur & Bode, 2014). According to Liu (2013), this meaning problem can be solved by bringing together separate modes of analysis in the form of combinations of “close” and “distant” readings, that is, increasing the understanding of the literature by studying particular texts in combination with aggregation and analysis of large amounts of data. In other related fields, such as media studies and media and communication studies, digital media and methods include discussing and dealing with ontology and epistemology in relation to public issues and difficulties for research (Hutchinson, 2016). Here, big and small data span from being part of a general critical discussion (Fuchs, 2017) and the logic of datafication (van Dijck, 2014) to methodological aspects of relating qualitative to quantitative methods (Mills, 2017; Schröder et al., 2012).

Besides these approaches, big data analysis in particular involves more computational methods, like artificial intelligence, Bayesian statistics and similar clustering techniques. In these areas the meaning problem from a sociological understandings means that the major methodological problem is that people cannot really say what they mean and how they make sense of things (Kåhre, 2009). The major methodological problem then, according to Kåhre, is that people cannot really say what they mean and how they make sense of things. They can only explain how they feel they have come to the knowledge,

because we humans cannot experience these processes other than through our brains. This relates to the way in which audience research in more contemporary digital media studies is dealing with the meaning problem. Here, one issue regards sense making of self-perceived media use in relation to insights of observed digital traces of the same use, underpinned by epistemological and ontological differences regarding how to undertake such studies online (e.g., Jensen, 2012; Rogers, 2013).

Thus, the meaning problem of big and small data asks important questions concerning how this configures science (including interpretation). The inference problem – how to infer meaning from data – is, however, present in all research readings, interpretations and arguments, not only in studies of critical big data. Here, there seems to be a meaning problem related to how data as combinations of methods and interpretations can be approached. This meaning problem is not new, but, by sketching a broad description of various research fields' approaches to method triangulation, a foundation can be made for what this could mean in digital method triangulation.

Overview of method triangulation

An onlife perspective presupposes a contextual understanding. When approaching the meaning problem of big and small data as combinations of methods, an important contextual understanding is the history of the development and use of the combination of methods in terms of method triangulation. The history of method triangulation shows how triangulation intends to combine two or more aspects of research to strengthen the design of a study and to increase the possibilities of interpretation of the results, combining data sources, researchers, theories, analytics, methods and creating a form of cross-verification (Arthur & Bode, 2014; Berry, 2012; Denzin, 1970; Fetveit, 2000; Geertz, 1973; Jankowski & Wester, 2002; Jick, 1979; Morley & Silverstone, 2002; Thurmond, 2001). Differences in events, situations, times, places and people reveal the atypical or show similar patterns, thereby increasing the validity. In the present study, we focus on the triangulation of *methods*, in particular what this could mean in an onlife approach conceptualized as digital method triangulation.

Triangulation is similar to, and can imply, the concept of mixed-method or multi-method research. Both are forms of “diverse testing” (Miller & Gatta, 2006: 597), but triangulation requires the framing of an issue to confirm it – compare for example trigonometry, in which triangulation is used to identify the precise location of a point – whereas the mixed- or multi-method concept is wider in scope (Fetters & Molina-Azorin, 2017). Two types of method triangulation are intra-method triangulation and intermediate or crossing method triangulation. From a historical standpoint, the former means that the researcher uses at least two data collection procedures from the same (qualitative *or* quantitative) approach, compared with the latter, in which both approaches are used in the same study (Thurmond, 2001). However, Denzin (2010) claimed that triangulation traditionally has been almost synonymous with qualitative research, in which the researcher is very close to the data. Thus, the risk of subjectivity is high but can be reduced with the use of triangulation (see Mackey & Grass, 2005). Moreover, triangulation has been argued to have been more common in social sciences than in the humanities, mainly because the concept of methodology has been considered to be less relevant to humanistic research (Fetveit, 2000). One exception is the field of digital humanities,

in which a large part of the efforts to explore the potential of digitalization – both as a phenomenon to investigate and as method development – is taking place today (Arthur & Bode, 2014; Berry, 2012). Both social sciences and the humanities have, however, also previously recognized the possibilities of triangulation in the use of the quantitative. Triangulation has been used to assess comparative research, for example in comparisons of different countries and in audience studies of media use. Ethnographic methods are also based on, and depend on, different triangulation techniques, mainly qualitative methods, in which more data sources, or rather research materials, reduce the risk of the study becoming method dependent (see Morley & Silverstone, 2002).

Validation in particular has been raised as an argument for triangulation (Jick, 1979). The potential weaknesses that can occur from using a single method can be overcome and the research can become more independent from researcher bias (Denzin, 1970). Triangulation can also stimulate the innovative use of known methods for unexpected dimensions within the subject; with appropriate theoretical and meta-theoretical reflections, it can provide more certainty in conclusions, above all in qualitative studies, and assist in constructing a more comprehensive perspective on specific analyses (Geertz, 1973; Jick, 1979). Triangulation is thus considered valuable in the development of methodology as well as theory in both humanities and social science (e.g., Jankowski & Wester, 2002). The challenge of triangulation, independent of the epistemological standpoint, is to find the right combination of methods, data and researchers and the best way to compare and analyse them. Triangulation also often takes more time and is more complicated to use than just one method.

For a long time, however, the epistemological differences between qualitative and quantitative standpoints on how insights can and should be achieved have prevented any cross-fertilization between the two approaches (Patriarche et al., 2014). However, according to Patriarche and colleagues, the methodological discussion today is more pragmatic and less ideological. Researchers have started to think about how to use the respective strength to gain further insights (Schröder et al., 2012) and how combinations can be developed so that these strengths can be utilized fully (Jankowski & Wester, 2002). Nonetheless, different ontological and epistemological standpoints also exist to various extents in the digital realm.

Digital method development has been suggested to be based on two different perspectives: offline and online grounded. The first perspective sees digital method development as reuse or digitalization of traditional or conventional methods (also referred to as “virtual methods”) and the second as digitally anchored or created methods. Two influential researchers in this area are Klaus Bruhn Jensen and Richard Rogers, advocating the respective perspective.

Jensen (2012) argued that methods, like media themselves, are “reused” or “recycled” as a response to digital media development, especially methods for data collection and analysis. Digital media involve ongoing documentation of various amounts of data in which the digital system can be seen as the actual method. At the same time, the fact that the documentation of data is spread and locally embedded poses a challenge. Jensen therefore distinguished between “found” and “made” data. Found data are automatically generated in the digital and are a form of secondary data. Made data are instead contextualization of data in a meaningful way and are often primary data. According to Jensen, a methodological challenge is to understand and reconcile these extremes; for

example, ethnographic methodology copes with this challenge through digital and virtual ethnography (e.g., Hine, 2000). Rather than replacing made data, big (found) data, for example, can function as a complement and an addition to the multi-methodological evidence box in audience studies. Jensen (2014) argued that we need all the methods that we have, offline grounded as well as online grounded, to capture and understand users in an increasingly complex onlife.

For Rogers (2013), digital methods “follow the medium”, which means methods embedded in the media processing online data, for example how search engines, such as Google, handle links and clicks and how applications, like Facebook, handle interactions, for example linking and sharing. Digital methods are used to collect and analyse hyperlinks, tags, search engine results, archived websites and other digital objects. Digital methods also try to embrace a social approach, for example when a search engine is used to study social change, such as *Google Flu Trends* (Rogers, 2015). A key question in relation to the digital-method context is how we can learn from search engines and recommendation systems about cultural change and social conditions regarding online dynamics. Thus, digital methods include a socio-material approach to method development, thinking with devices and objects for researching cultural change and social conditions. Rogers was critical of the fact that online data in many cases today methodologically seem to need to be rooted in offline data and that “offline becomes the control of which online quality is measured” (2015: 3). Digital methods thus raise the question of the prerequisites for online anchoring: under which conditions can for example media studies be based on or in online data?

Hence, the challenges of and opportunities for digital method triangulation lie in the historical basis of triangulation, the researcher’s methodological standpoint, the nature of the digital and the increased possible combinations for studying the onlife. The concept of digital method triangulation directs particular attention to methodological development in the field of social science and humanities, incorporating the way in which digital structuring matters as well as the conditions and meanings of media studies of the onlife.

Example 1: Digital method triangulation through digital focus groups

Focus groups are one of the most widely used methods within small data studies and qualitative methods (Carey & Asbury, 2012; Halkier, 2010; Nyumba et al., 2018). Digital focus groups as an example of triangulation in small data studies provide insights into digital method triangulation possibilities, such as what an online context brings to the focus group method.

Focus groups are a data collection method whereby an interviewer interviews several people at the same time in a focused group discussion to obtain qualitative insights into what and how people think of a particular theme. The participants are often relatively homogeneous and unknown to each other. Focus groups are usually conducted as separate, several and consecutively made focus groups (Barbour & Kitzinger, 1999; Krueger, 1994; Liamputtong, 2011).

Digital media development has stimulated the development of so-called *virtual focus groups* (Stewart & Shamdasani, 2015) and *online focus groups* (Stewart & Williams, 2005). The following benefits have been suggested: gathering groups of people who would otherwise fail to meet and reaching particularly young people; demonstrating the

role of the moderator or the absence of a role; being suitable for raising and discussing sensitive issues; gaining more transparency in group interaction; and creating other forms of information provision and documentation at the same time than is the case with offline focus groups (Stewart & Shamdasani, 2015; Williams et al., 2012).

Digital focus groups in the context of triangulation can be seen as potential combinations of these benefits in relation to *location*, *time* and the online aspect that also involves combinations of *data*.

Location means the potential for gathering data that are space and location independent. In research using digital focus groups, the location aspects have been explored and made use of to gather groups of people who would otherwise fail to meet. Turney and Pocknee (2005) created online discussion forums for focus groups and found that they strengthened the ability to extend to hard-to-reach populations that were otherwise scattered. Lijadi and van Schalkwyk (2015) came to realize that Facebook focus groups could connect the global and the local as well as creating close interaction in a small context of interest groups.

Gathering data through virtual/online focus groups is mainly approached as gathering people in relation to time as synchronous: simultaneous. Then, focus groups are used to engage people collectively in a focused conversation at the same time, which reuses the offline characteristics of the focus group. This is not surprising, because it is a fundamental characteristic of an offline focus group. Reusing offline methods has a tendency to bring fundamental characteristics into an online context (Rogers, 2013).

However, focus groups can also benefit from making use of the online context potential. Concerning time, this means that focus groups can take place asynchronously: on different occasions, ongoing or even more dispersed in time. With this rather simple temporal possibility, several new aspects of the focus group can be discovered and combined. Williams and colleagues (2012) argued that an asynchronous focus group is suitable for sensitive aspects in that it allows participants to construct the stories that they want and gives the researcher more time to consider participants' responses, which can reduce the risk of misinterpretation. Hence, the asynchronous focus group seems to open several aspects and interpretative acts that can be considered as valuable to see and to use by themselves or in combination.

A study could combine synchronous analogue and digital focus groups as well as combining various asynchronous digital focus groups. This could enhance the benefits of face-to-face research by including multiple participants' experiences in new forms of narratives and providing other interpretative prerequisites. In larger focus groups enabled online, several different interest groups can be formed and possibly also become several other focus groups. Hence, a study trying to combine a digital focus group with an offline focus group also reveals variations in the researcher's moderator role and function. How the role of the moderator and moderation changes in digital contexts needs to be explored more (Williams et al., 2012). A question furthering this knowledge production is: What does it mean to handle multiple conversations and stories simultaneously, when the conversation does not need to be coordinated with something completely predetermined and in common?

Online platforms' technical attributions can also be understood as a form of co-moderator. This directs attention to how digital focus groups can include an understanding of platforms' characteristics and conditions. Social media as spaces for digital focus

groups also differ, not least in their perception of them, from more closed forums and other forums initiated by researchers. Digital method triangulation can very well mean understanding a video conferencing focus group in relation to, for example, a Facebook group, in which the multimodal plays a role with various metadata possibilities. The data from digital focus groups are written on the screen in text, in the same way as interview questions and answers, but they can also be pictures and sounds. Available found data also include when posts are made and/or shared and by whom, allowing timelines or a social networking analysis of who responds to whom and interacts most and least, respectively. Group dynamics can thus be a major part of the analysis of digital focus groups. Digital focus groups are organized in a way that makes them primarily understood as made data. However, digital focus groups can be combined with found data understood as the “automatic” online traces of data and metadata, like time stamps and geolocations.

If using the digital focus group method with social media platforms, an important question concerns bias: whether social media users are representative selections and how to handle possible skewness by not primarily having generalization claims. Focus groups often stand as representatives of certain target and interest groups. However, studying fan groups of popular cultural phenomena as online focus groups does not necessarily mean reaching and gaining the selection of the anticipated demographics. Online fan groups make visible the fact that the main group-defining attribution is the fan object, like a certain popstar, not the demographics. At the same time, other online fan communities might attract only certain demographics. Even though many people use social media, a lot of people do not. The purpose and the context thus determine the representativeness and generalizability, and this relates to the aim of the researcher. Aspects that remain and are relevant to selection are linked to language and online presence, including the habits of managing the platform used and the attitude towards and views of sharing online. The latter is also connected to ethics, an important part of making sense of big and small data and digital method triangulation.

The research ethics for digital method triangulation need an approach that is sensitive to what has been conceptualized as networked privacy. The ethical approach of networked privacy is based on how social media change the practices of information exchange and visibility and show the shift from privacy perceptions to something in context and a network, which makes integrity an ongoing active internship (Marwick & boyd, 2014). What is considered to be private and informed consent is closely linked to the perception of the people who contribute the research material or data. This contradicts the idea of forms of triangulation as a “snowball method”: starting the data collection and then detecting more data being studied but without informed consent about these particular data. This also involves challenges related to data visualization, which does not gather more data but transforms it so that it can be perceived as new data.

Digital focus groups as digital method triangulation involve the question of where users/people/audiences are and what these locations as conversation groups (more or less moderated) can say about people and their meaning making.

From method triangulation, we know that surveys can be combined with interviews in two ways to strengthen the interpretation: 1) to use interviews to identify qualities and to use a survey to generalize how much/how many these qualities are; 2) to use a survey first to find individuals and groups that are particularly interesting for gaining deeper

knowledge by interviewing them. From digital method triangulation, we can make use of not only self-assessed, perceived and expressed data/information but also the online traces of the focus group. These online traces are not (only) traffic data, as in the method example below (comparing what is said and what is done). Instead, a primary advantage of the online traces in a digital focus group is that the study is strengthened by including data that can give a perspective on what influences focus group talk.

Example 2: Triangulation of big(ger) data and virtual methods: Comparing measured internet use with self-perceived use

Combinations of traffic data measurements (see the definition below) of users' internet behaviour with surveys and diaries exemplify in several ways what digital method triangulation can entail. It includes the relationship between digital (traffic measurements) and virtual methods (surveys and diaries), big (traffic measurements) and small data (surveys but in particular diaries), observed data (traffic measurements) and self-perceived data (surveys and diaries) and qualitative (diaries) and quantitative (traffic measurements and surveys) approaches.

Tracking or capturing user behaviour online has long been technically possible and has been a growing theme in the research literature since the 1970s (Smith et al., 2011). However, measuring technology has become better and cheaper – it is now possible to track other aspects than aggregated traffic at the IP level and to identify specific applications (Kihl et al., 2010). Knowledge of big data analysis has also grown, which has increased the opportunities for various actors in society to track users. Nevertheless, performing such measurements can still be quite complicated (e.g., Findahl et al., 2014; Taksa et al., 2009), since they require access to networks and measuring tools (hardware as well as software) and includes aspects such as handling large quantities of personal data, which can be challenging from the perspective of storage, analysis and legislation. In the realm of big data analysis, Manovich (2011: 10) thus talked about three data classes: “those who create data (both consciously and by leaving digital footprints), those who have the means to collect it, and those who have expertise to analyze it”. The second and especially the third group are still small.

Traffic measurements are a collective term referring to the practice of capturing the trail of data – often through IP addresses – that is created between client and server computers (uploads and downloads) on the internet every time a user's computer requests data from an internet actor (e.g., Kihl et al., 2010; Lagerstedt et al., 2012). Originally traffic measurements were used to adapt and optimize internet access networks to the requirements of customers from an engineering perspective (e.g., Kihl et al., 2010; Ladner, 2009). The measurements can – with different depths and details depending on how one measures and with which tools – include data such as audience interactions (e.g., exposure and time), social interactions (e.g., posts and likes) and platform interactions (ID, dates and tags) (cf. Giglietto et al., 2012), including activities that are unknown to the user as well as non-human-induced internet traffic. Thus, traffic measurements often entail collecting *all* the traffic on a broadband access network(s) but can be used to capture traffic from specific applications or protocols, similar or equal to the collection of, for example, web log files or hashtags used in social science (e.g., Menchen-Trevino & Karr, 2012; Vicente-Mariño, 2014). Although the method is more common in fields

like computer science and media technology, it can also provide great opportunities for studies in social science and the humanities (e.g., Manovich, 2012), particularly in combination with other methods.

Like other ways to track and capture behaviour data online, the great advantage of traffic measurements is the fact that actual audience behaviour is measured, reducing the influence of human bias compared with recalled behaviour. Since the measurements are carried out without any intermission from the observer, potential research biases can also be minimized (Findahl et al., 2014). Thus, the method provides not only data about the user, the platform or the application used but also an opportunity to understand aspects that are not directly observable (Giglietto et al., 2012). One advantage over the other tracking methods is that the traffic measurement tool cannot be disabled by the subject (cf. Scharrow, 2016). However, the method also involves several limitations. As with other tracking methods, traffic measurements measure potential exposure, not actual attention – the application can be running despite no one being present or someone can be present without paying attention (cf. Greenberg et al., 2005; Lagerstedt et al., 2012). Moreover, even though the current systems can measure all the users' internet traffic and the traffic per IP number, they still often work per household and not per individual or specific user; thus, it is not possible to know exactly who is generating the data without making further inquiries. Internet traffic also gives rise to activities that are not induced by the user, such as automatic updates. Thus, the total traffic volume recorded can exceed the user's actual use. Moreover, some applications are more traffic heavy, or compressed, than others, which can result in a skewed image of the internet traffic that is actually generated.

Due to the challenges of traffic measurement of user behaviour, mixed methods and triangulation have been recommended, but they are not yet common (see Findahl et al., 2014; Giglietto et al., 2012; Lagerstedt et al., 2012). Some studies have combined observed online data with other methods (e.g., Araujo et al., 2017; Scharrow, 2016), but few have used triangulation. Findahl and colleagues (2014) and Lagerstedt and colleagues (2012) studied internet audience behaviour through triangulation of traffic data measurements and surveys and diaries, adopting a meta-perspective. Their results showed good coherence between behaviour and experience in terms of activity and perception of time; people who were active internet users according to the traffic measurements were also active according to the surveys and diaries and vice versa, and activities stated as common through the conventional or virtual methods were also detected in the traffic measurements. However, the studies found that the coherence between the methods was dependent on the type of activity. It was higher for the use of websites than for streaming media. One explanation was that streaming media, audio in particular, can be a secondary activity and therefore it is more difficult to exactly assess when it occurred; furthermore, the questions in the questionnaires were not formulated in a way that captured this behaviour well. However, the results showed that the compliance increased if the activity (e.g., listening to music) could be linked to a specific service (e.g., Spotify) (see also Scharrow, 2016). In both studies, the use of triangulation showed that neither the questionnaires nor (especially) the diaries gave a complete picture of the users' actual internet behaviour (see also Greenberg et al., 2005; Scharrow, 2016). One possible reason suggested was that internet use often includes multiple activities that are simultaneously ongoing (multitasking) and another was that the term "surfing the

internet” can be interpreted as, and mean, many different things. Moreover, the results confirmed a weakness of tracking as a method for measuring behaviour; the traffic measurements of the users’ behaviour exceeded their perceived use marked in the diaries. The reason for this is that a user’s internet traffic consists of applications that run without the user being present (for example at night) and traffic that is machine induced. Thus, traffic measurements measure a user’s internet traffic correctly, but, since not all traffic is human induced, the user’s behaviour is not accurately displayed.

The traffic measurements showed that the virtual methods did not give an incorrect but an incomplete picture of users’ behaviour online (see also Araujo et al., 2017; Scharkow, 2016); surveys only produce answers for the questions asked, whereas data measurements can provide more extensive information, especially for activities that change over time. The diary method provides a greater ability to capture change, but only the ones that the users find to be the most important, whereas data measurements capture all activities (Findahl et al., 2014). However, while surveys and diaries are associated with weakness prone to human error, such as memory problems, time estimations and untrue information (Dienlin & Trepte, 2015; Findahl et al., 2014), they are especially valid in capturing the active role of the audience in mediated processes: attitudes, beliefs and needs (Findahl et al., 2014; Vicente-Mariño, 2014). Traffic measurements, on the other hand, include redundant information, as they also capture non-human-induced activities. Thus, the triangulation showed that the methods were mutually confirmatory as well as complementary, providing a more complete picture of the complex behaviour that occurs onlife than each single method would have offered. Triangulation thus compensates for the weaknesses of each method by counterbalancing their strengths (cf. Denzin, 1970). Digital method triangulation, however, still requires development to enable the methods to be more directly comparable.

Discussion

In this article, we have raised the issue of how digital method triangulation can be an appropriate and at the same time ongoing development of combinations that involve forms of digital anchoring, which is unusual in, for example, other digitized methods, such as virtual ethnography.

What onlife can mean for digital method triangulation concerns how to integrate the online and offline realms when combining various field sites. It is the awareness of choosing the right method for answering the desired question and for being aware of the new research questions that digital media can generate.

Digital focus groups invite combinations, implying research questions from both the “offline context” and the “online context”. Place-located focus groups, in which participants are often relatively homogeneous and unknown to each other, indicate a desired research question to be able to gain qualitative insights into what and how people think of a particular theme. In addition, in a more digital context, desired research questions are related to an understanding of how to gather online groups or people, the moderator role, sensitive issues, group interaction and forms of data and documentation (traces). New research questions can mean more involvement in communities of interest rather than location-dependent interests to study the historical development of a community of interest. This means for example asking questions like the following: Which location-

scattered communities of interest has the research field not studied before? What kind of research questions arise in close interaction between participants rather than being part of a strongly moderated discussion? How can larger focus groups enabled online become smaller and develop into other focus groups and why does this happen? As a consequence, onlife approaches are better taken into account in that they articulate how the digital focus group (participants and moderator) is embedded in everyday onlife practices and how the offline and online methodological understandings relate to each other.

Traffic measurements combined with surveys and diaries are a valuable form of triangulation for answering questions in the digital related to internet user behaviour. The example showed that the digital method triangulation compensated for the weaknesses of each method by counterbalancing the strengths of the others. By being mutually confirmatory as well as complementary, triangulation provided a more complete answer regarding users' onlife behaviour – an answer that would not have been possible to achieve with the digital method alone. However, triangulation raised new questions on the meaning making of different internet phenomena and how to capture multiple and interconnected modes of communication. For example, what does the action of “surfing” mean to different people? How can multiple activities that are simultaneously ongoing be understood more accurately? How can machine- and human-induced activities be separated more correctly? To answer these questions, methodological adjustments or new combinations may be required.

Moreover, a challenge in big data studies is the often-taken-for-granted causal significance of volume and number: for example, if an application generates large volumes of data, or if a link is shared numerous times, it is considered to be a sign of importance. This is particularly evident in high-rank approaches in which link shares and interactions are seen as signs of power and relevance. While not evident in the examples of the article, this is where a combination of methods has also been suggested to be important in sensitizing taken-for-granted causalities that otherwise would not have been made visible.

Thus, digital method triangulation seems to provide valuable knowledge for method development. This article provides some reference points for navigation. A primary understanding is that combinations are more than just the use of both qualitative and quantitative, or online and offline, methods. Understanding includes knowledge of the possibilities and challenges of combinations that exist within the digital. It is an awareness of what is combined – method, data, sources or researchers – more or less under the same theoretical umbrella, where one position is the view of “in the medium”. For the researcher, this requires knowing and selecting combinations in a basic way as well as understanding the conditions and features of the digital to address its diversity. However, the following question remains: How can researchers determine which methods can be combined and how? We argue that this is normative and depends on the ontological and epistemological experience and goal of the researcher. Jensen (2014) claimed that future studies of users have to acknowledge media as constituents of technical configurations. To study these aspects, research methodologies need to consider the diversity of data – large or small, found and made, qualitative and quantitative; users are everywhere, both online and offline. While these onlife contexts may be inextricably interwoven, they can also be increasingly complex; thus, multiple tools and combinations are required to make sense of them. In these situations, concepts such as digital method triangulation can be a possible way forward.

Digital method triangulation for making sense of small and big data has several implications: 1) it can stimulate innovative use of known methods for unexpected dimensions within the topic; 2) with appropriate theoretical and meta-theoretical reflections, it can provide more certainty in conclusions; and 3) it can assist in constructing a more comprehensive perspective on specific analyses. Hence, digital method triangulation can help to decrease the fixation on the measurability of the digital as well as not having to renounce the offline in making the online visible while at the same time being able to take advantage of the new possibilities that the digital offers. To catch all the nuances that current media research requires, one has to adopt a wide approach. Digital method triangulation offers an opportunity to catch the small as well as the big in current media use.

The three implications presented above not only make the old validation aim of triangulation explicit. Digital method triangulation involves paying more direct attention to method development, in which known methods are challenged in redirections of new use. Users as part of both online and offline contexts particularly articulate dimensions of meaning making, whereby the former self-perceived meanings and causal links become more of a challenging development process, searching for contextual strengths and avoiding weaknesses. This can be expressed more as an experimental process in which validation is of interest but not pre-set or even pre-known. In this process, it seems that method development is needed more than ever to occur in a dialogue with theory. Validity or certainty in conclusions – generalizable and transferrable depending on the ontological and epistemological standpoints of the study – is created with, and through, theoretical and meta-theoretical reflections. Making sense of big and small data cannot only rely on combinations of methods. There is a need for meaning making of the combinations, in which theory is crucial. Inferring meaning from the digital traces to the incentives, motives and needs of humans cannot only be undertaken through comparisons of different data obtained from different methods. Digital method triangulation as a way of making sense of big and small data shows how the specific analysis of a study, in a particular context, is foci. A comprehensive perspective on specific analyses is the key.

Conclusion

The increased awareness of how media scholars can develop their knowledge of using method triangulation for meaningful inferences and interpretation of big and small data in the onlife context can produce many implications. In this article, we have shown how digital method triangulation facilitates dialogue between virtual and digital methods, which appears to be crucial to capture the complexities of the onlife. The validity of combining methods has a history and can benefit from learning from triangulation research. The need for theoretical reflection can benefit from including critical big data studies. From this perspective, a particular conclusion of that critical interrogation of assumptions and biases is that not only data but also the methods and the combination of methods in configuring science come from somewhere.

References

- Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired*, 23 [online]. Retrieved from <https://www.wired.com/2008/06/pb-theory/> [accessed 2019, February 4].
- Araujo, T., Wonneberger, A., Neijens, P. & de Vreese, C. (2017). How much time do you spend online? Understanding and improving the accuracy of self-reported measures of internet use. *Communication Methods and Measures*, 11(3): 173-190.
- Arthur, P. L. & Bode, K. (2014). *Advancing digital humanities: Research, methods, theories*. London: Palgrave Macmillan.
- Bailenson, J. N. (2012). Doppelgangers: A new form of self. *The Psychologist*, 25(1): 36-39.
- Barbour, R. & Kitzinger, J. (eds.) (1999). *Developing focus group research: Politics, theory, and practice*. London: Sage.
- Berry, D. (2012). *Understanding digital humanities*. Basingstoke: Palgrave Macmillan.
- Boyd, D. & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5): 662-679.
- Carey, M. A. & Asbury, J.-E. (2012). *Focus group research*. New York: Routledge.
- Denzin, N. K. (1970). *The research act: A theoretical introduction to sociological methods*. Chicago: Aldine.
- Denzin, N. K. (2010). Moments, mixed methods, and paradigm dialogs. *Qualitative Inquiry*, 16: 419-427.
- Dienlin, T. & Treppe, S. (2015). Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors. *European Journal of Social Psychology*, 45(3): 285-297.
- Dourish, P. & Button, G. (1998). On "technomethodology": Foundational relationships between ethnomethodology and system design. *Human-Computer Interaction*, 13(4): 395-432.
- Fetters, M. D. & Molina-Azorin, J. F. (2017). The Journal of Mixed Methods Research starts a new decade: Principles for bringing in the new and divesting of the old language of the field. *Journal of Mixed Methods Research*, 11(1): 3-10.
- Fetveit, A. (2000). Den trojanske hest: Om metodebegreper marginalisering av humanistisk medieforskning [The Trojan Horse: About the method concepts marginalization of humanistic media research]. *Norsk Medietidskrift*, 2: 5-27.
- Findahl, O., Lagerstedt, C. & Aurelius, A. (2014). Triangulation as a way to validate and deepen the knowledge about user behavior: A comparison between questionnaires, diaries and traffic measurements. In G. Patriarche, H. Bilandzic, J. Linaa Jensen & J. Jurisic (eds.), *Audience research methodologies: Between innovation and consolidation* (pp. 54-72). London: Routledge.
- Fuchs, C. (2017). *Social media: A critical introduction*. London: Sage.
- Geertz, C. (1973). *The interpretation of cultures*. New York, NY: Basic Books.
- Giglietto, F., Rossi, L. & Bennato, D. (2012). The open laboratory: Limits and possibilities of using Facebook, Twitter and YouTube as a research data source. *Journal of Technology in Human Services*, 30(3-4): 145-159.
- Greenberg, B. S., Eastin, M. S., Skalski, P., Cooper, L., Levy, M. & Lachlan, K. (2005). Comparing survey and diary measures of internet and traditional media use. *Communication Reports*, 18(1): 1-8.
- Halkier, B. (2010). Focus groups as social enactments: Integrating interaction and content in the analysis of focus group data. *Qualitative Research*, 10(71): 71-89.
- Hine, C. (2000). *Virtual ethnography*. London: Sage Publications.
- Hutchinson, J. (2016). An introduction to digital media research methods: How to research and the implications of new media data. *Communication Research and Practice*, 2(1): 1-6.
- Iliadis, A. & Russo, F. (2016). Critical data studies: An introduction. *Big Data & Society* [online 2017].
- Jankowski, N. & Wester, F. (1991 [2002]). The qualitative tradition in social science inquiry: Contributions to mass communication research. In K. B. Jensen & N. Jankowski (eds.), *A handbook of qualitative methodologies for mass communication research* (pp. 44-74). London: Routledge.
- Jensen, K. B. (2012). Lost, found, and made: Qualitative data in the study of three-step flows of communication. In I. Volkmer (ed.), *The handbook of global media research* (pp. 435-450). Hoboken, NJ: Wiley-Blackwell.
- Jensen, K. B. (2014). Audiences, audiences, everywhere: Measured, interpreted and imagined. In G. Patriarche, H. Bilandzic, J. Linaa Jensen & J. Jurisic (eds.), *Audience research methodologies: Between innovation and consolidation* (pp. 227-240). London: Routledge.
- Jick, T. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, 24(4): 602-611.
- Kåhre, P. (2009). *På AI-teknikens axlar: om kunskapssociologin och stark artificiell intelligens* [On the shoulders of AI technology: about knowledge sociology and strong artificial intelligence]. Lund Dissertations in Sociology, 1102-4712; 87. Lund: Department of Sociology, Lund University.
- Kihl, M., Lagerstedt, C., Aurelius, A. & Ödling, P. (2010). Traffic analysis and characterization of Internet user behavior. Paper presented at the conference *International congress on ultra-modern telecommunications and control systems and workshop* (ICUMT), 2010, October 18-20, Moscow.

Exploring the Meaning Problem of Big and Small Data Through Digital Method Triangulation

- Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society* 2014: 1-12.
- Kitchin, R. & Lauriault, T. (2014). *Towards critical data studies: Charting and unpacking data assemblages and their work*. The Programmable City Working Paper 2. Programmable City, Social Science Research Network.
- Kitchin, R. & Lauriault, T. (2015). Small data in the era of big data. *GeoJournal*, 80: 463-475.
- Krueger, R. (1994). *Focus groups: A practical guide for applied research*. London: Sage.
- Ladner, S. (2009). Watching the web: An ontological and epistemological critique of web-traffic measurement. In B. J. Jansen, A. Spink & I. Taksa (eds.), *Handbook of research on web log analysis* (pp. 504-520). New York, NY: Information Science Reference.
- Lagerstedt, C., Findahl, O., Aurelius, A., Pathirana, H. & Popp Larsen, C. (2012). Understanding Internet user behavior: Towards a unified methodology. *International Journal of Advances in Telecommunications*, 5(3-4): 153-163.
- Liamputtong, P. (2011). *Focus group methodology: Principles and practice*. Sage: London.
- Lijadi, A. & van Schalkwyk, G. (2015). Online Facebook focus group research of hard-to-reach participants. *International Journal of Qualitative Methods*, 14(5): 1-9.
- Liu, A. (2013). The meaning of the digital humanities. *PMLA*, 128: 409-422.
- Lupton, D. (2014). *Digital sociology*. London: Routledge.
- Mackey, A. & Grass, S. M. (2005). *Second language research: Methodology and design*. London: Lawrence Erlbaum Assoc.
- Mahrt, M. & Sharkow, M. (2013). The value of big data in digital media research. *Journal of Broadcasting & Electronic Media*, 57(1): 20-33.
- Manovich, L. (2011). *Trending: The promises and the challenges of big social data*. Retrieved from <http://manovich.net/index.php/projects/trending-the-promises-and-the-challenges-of-big-social-data> [accessed 2019, February 4].
- Manovich, L. (2012). Trending: The promises and the challenges of big social data. In M. K. Gold (ed.), *Debates in the digital humanities* (pp. 460-475). Minneapolis, MN: University of Minnesota Press.
- Marres, N. (2017). *Digital sociology: The reinvention of social research*. Cambridge: Polity Press.
- Marwick, A. E. & Boyd, D. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*, 16(7): 1051-1067.
- Mehl, M. & Gill, A. (2010). Automatic text analysis. In S. Gosling & J. Johnson (eds.), *Advanced methods for conducting online behavioral research* (pp. 109-127). Washington, DC: American Psychological Association.
- Menchen-Trevino, E. & Karr, C. (2012). Researching real-world web use with Roxy: Collecting observational web data with informed consent. *Journal of Information Technology & Politics*, 9(3): 254-268.
- Miller, S. & Gatta, J. (2006). The use of mixed methods models and designs in the human sciences: Problems and prospects. *Quality & Quantity*, 40(4): 595-610.
- Mills, K. (2017). What are the threats and potentials of big data for qualitative research? *Qualitative Research* 18(6): 591-603.
- Morley, D. & Silverstone, R. (2002 [1991]). Media audiences, communication and context: Ethnographic perspectives on the media audience. In K. B. Jensen & N. Jankowski (eds.), *A handbook of qualitative methodologies for mass communication research* (pp. 149-162). London: Routledge.
- Nyumba, T. O., Wilson, K., Derrick, C. J. & Mukherjee, N. (2018). The use of focus group discussion methodology: Insights from two decades of application in conservation. *Methods in Ecology and Evolution*, 9: 20-32.
- Patriarache, G., Bilandzic, H., Linaa Jensen, J. & Jurisic, J. (eds.) (2014). *Audience research methodologies: Between innovation and consolidation*. London: Routledge.
- Rogers, R. (2013). *Digital methods*. Cambridge, MA: The MIT Press.
- Rogers, R. (2015). Digital methods for web research. In R. Scott & S. Kosslyn (eds.), *Emerging trends in the social and behavioral sciences* (pp. 1-22). Hoboken, NJ: John Wiley & Sons.
- Scharkow, M. (2016). The accuracy of self-reported Internet use: A validation study using client log data. *Communication Methods and Measures*, 10(1): 13-27.
- Schröder, K. C., Hasebrink, U., Hölig, S. & Barker, M. (2012). Introduction: Exploring the methodological synergies of multimethod audience research. *Participation Journal of Audience and Reception Studies*, 9(2): 643-647.
- Simon, J. & Ess, C. (2015). The ONLIFE initiative: A concept reengineering exercise. *Philosophy and Technology*, 28(1): 157-162.
- Smith, J. H., Dinev, T. & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, 35(4): 989-1015.
- Stewart, D. W. & Shamdasani, P. N. (2015). *Focus groups: Theory and practice* (3rd ed.). Los Angeles, CA: Sage.
- Stewart, K. & Williams, M. (2005). Researching online populations: The use of online focus groups for social research. *Qualitative Research*, 5: 395-416.

- Taksa, I., Spink, A. & Jansen, B. J. (2009). Web log analysis: Diversity of research methodologies. In B. J. Jansen, A. Spink & I. Taksa (eds.), *Handbook of research on web log analysis* (pp. 504-520). New York, NY: Information Science Reference.
- Thurmond, V. A. (2001). The point of triangulation. *Journal of Nursing Scholarship*, 33(3): 253-258.
- Turney, L. & Pocknee, C. (2005). Virtual focus groups: New frontiers in research. *International Journal of Qualitative Methods*, 4(2): 32-43.
- Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, 12(2): 197-208.
- Vicente-Mariño, M. (2014). Audience research methods: Facing the challenge of transforming audiences. In G. Patriarcho, H. Bilandzic, J. Linaa Jensen & J. Jelena (eds.), *Audience research methodologies: Between innovation and consolidation* (pp. 37-53). London: Routledge.
- Williams, S., Giatsi Clausen, M., Robertson, A., Peacock, S. & McPherson, K. (2012). Methodological reflections on the use of asynchronous online focus groups in health research. *International Journal of Qualitative Methods*, 11(4): 368-383.