

Netnography and a Summative Content Analysis Approach to Market Research

Alan SHAW

Leeds Beckett University
alan.shaw@leedsbeckett.ac.uk

Abstract

With organisations like Facebook restricting how their application programming interface (API) can be used and scholars questioning the legality and ethics of web scrapping (i.e., the use of technology in the automatic extraction of data from the Web) more discussions around a qualitative Netnographic approach is needed. This paper addresses these issues by reflecting on the application of a passive summative content analysis method to Netnography and how it can be used in marketing research. It focuses on the rollout of smart meters (meters that allow consumers and service providers to monitor power consumption), which the UK Government has now delayed because of a poor uptake. As such, it contributes to the marketing domain's theory and knowledge and provides a possible set of solutions that the UK Government and energy providers could consider to increase engagement. The study starts by providing an overview of the literature within Netnography and its use as a qualitative methodology. It then demonstrates, step by step, how a summative content analysis approach can be applied to Netnography, using NVivo as the platform of analysis. The case study utilises Mumsnet (UK's biggest network for parents, with approximately 10 million unique visitors and 100 million-page views per month) as the forum for analysis. Threads over a six-month period were considered. The key themes identified can be explained as: smart meters were not transferable between energy providers; users were concerned about being hacked; the connecting signals did not always work; and such meters were not compulsory. The study demonstrates how effective and efficient Netnography can be in market research. It also provides some clear guidance on how copyright issues should be addressed.

Keywords: Netnography, Summative content analysis, Market Research, Smart meters.

JEL classification: C90, D83, M30, O30.

1. Introduction

Netnography is an adapted ethnographic approach that uses archival and real-time data from all internet enabled technologies (Kozinets, 2015). It can be quantitative (see Belz and Baumbach, 2010; Hardy, 2017) or qualitative (see Bartl et al., 2016; Elliot et al., 2005), although Kozinets (2015), who arguably is the 'father' of Netnography, only positions it as a qualitative approach. There are a variety of data capture processes used in Netnography, Reid and Duffy (2018) acknowledge that a platform's application programming interface (API), which is a piece of software that lets two applications talk to each other (Lauret, 2019), is the ideal mechanism for providing researchers with vast amounts of data to analyse. Venturini and Rogers (2019) have since identified that the data breach involving Cambridge Analytica and Facebook in 2014 has resulted in Facebook and a number of similar organisations restricting access to their APIs. This restriction makes the data capture process harder, so Venturini and Rogers (2019) have asked scholars to look at other means of gaining digital data. A possible alternative is to employ web scraping techniques (i.e., the use of technology in the automatic extraction of data from the Web), but scholars like Krotov and Silva (2018) question the legality and ethics of such a process. This dilemma could be a reason why Costello et al. (2017) posits that the Netnographic process is poorly understood. As such, this study will critically evaluate the issues relating to data capture and analysis in Netnography, using an adapted summative content analysis approach to demonstrate a possible solution. For illustrative purposes, a case study focusing on the rollout of smart meters (meters that allow consumers and service

providers to monitor power consumption) was considered. The author believes that this study contributes to the theory and knowledge within the marketing and research process domains because it addresses the gaps identified by Costello et al. (2017) and Venturini and Rogers (2019). In addition to this, the author believes that there is also a contribution to practice because the case study identified a number of possible solutions which could be used by the UK Government to address the poor uptake of smart meters by the general public (see OFGEM's (2019) open letter to key its stakeholders).

2. The Qualitative Approach to Netnography and Summative Content Analysis.

Data collection is one of the key elements within the Netnographic process (Kozinets, 2010), with the API protocol being the most efficient (Puschmann & Ausserhofer, 2017). As mentioned however, API restrictions and the legal/ethical issues associated with web-scraping have restricted the options open to researchers. The two remaining alternatives are to either analyse the data directly from the site under review or to copy and paste extracts onto another platform for dissection and manipulation. The first option is arguably more cumbersome and restrictive. It also means that the researcher is reliant on using the host platform for the back-up of their research data, this makes retrieval or reassessment difficult because more posts are likely to appear after the initial review. Copying and pasting the data onto an alternative platform would be a better solution, but this may lead to legal issues. Kozinets (1998) first raised copyright concerns in his initial assessments of Netnography. It was revisited again in 2014 (Kozinets et al., 2014) but the author's assessment of the literature associated with Netnography identified that the discussions relating to Netnography and copyright were very limited.

Copyright regulations are complex because they are country specific (Kozinets et al., 2014), this may explain why researchers have chosen to ignore it. There are however, mitigating arguments to counter possible copyright infringement claims. These mitigating arguments are linked to a doctrine known as "fair use" or "fair dealing". The USA uses *Section 107 of the Copyright Act* as its statutory framework for identifying if a copied item can be deemed as "fair use". The framework has a list of factors that must be considered before a final decision can be made. In the UK, the term "fair dealing" is used and is governed by *Sections 29 and 30 of the Copyright, Designs and Patents Act 1988*. From an academic research perspective, scholars who copy and paste data for Netnographic research could use the "fair use" or "fair dealing" argument to counter copyright infringement claims, that said the claim will still be open to interpretation. The argument of "fair use" and "fair dealing" are considered under common law jurisdictions (Band and Gerafi, 2013). Countries that adopt a civil law jurisdiction (i.e., Continental Europe) do not have a "fair dealing" copyright process. Europe sets out directives that are not legally binding but must be adapted (or transposed [The transposition of EU Directives is the process by which member states give effect to Directives within their own domestic legal system (LexisNexis, nd)]) by each member country. This means that each European member country may have a slightly different view. More concerning are Articles 15 and 17 (see Directive (EU) 2019/790), these will require authors to provide proof of licencing in the use of any copyright material, which could impact how editors view articles related to Netnography. There is also a wider debated in terms of how any "fair use" or "fair dealing" perceived materials will be viewed in all academic journals, but this goes beyond the scope of this article. Researchers from countries adopting a civil law jurisdiction will have to review their country's statute books before making a decision on their data captured method. A solution to this dilemma is to always approach the owner for permission, although it is the author's experience that engaging with the big social media networks like Facebook and Twitter is difficult.

Bartl et al. (2016) identified numerous qualitative approaches that could be applied to the Netnographic process. Kozinets (2020) believes that researchers should link these to either a passive approach (where there are no engagements with participants) or an active approach (where researchers would interact with participants). In contrast, Costello et al. (2017) advocate that users should only embrace the active approach. Lugosi et al. (2012) and Costello et al. (2017) argue that the passive application of Netnography should be reframed as ‘qualitative archival data research, and not Netnography. The author believes that such a proposition could be detrimental to the Netnographic domain, because even though active engagements could yield significant cocreation opportunities and/or identify richer sources of information, they do not address the ethical and procedural implications associated with such a stance. The Market Research Society [MRS] (2014) stipulate that all online research that has any participant engagement must gain informed consent, a view which is also shared by Keim-Malpass, et al. (2014). Heinonen and Medberg (2018) identified that the majority of Netnographic researchers used the passive approach because of the its ease of implementation. It confirms that passive qualitative Netnographic research remains an important methodology. In addition, Elliot et al. (2005) identified the passive approach to be less costly for market researchers to employ.

Vaismoradi et al. (2013) believe that scholars select qualitative methods because they want to gain a deeper understanding of a participant’s viewpoint for a given situation. Readers, however, will no doubt be aware that there is a plethora of options available to researchers employing qualitative techniques. As such, this study will restrict its overview to the thematic and content analysis approaches. The thematic analysis approach is widely used in the field of qualitative research but is rarely acknowledged (Braun and Clarke, 2006). Braun and Clarke (2006) also believe that some researchers mistakenly misclassify the type of researcher they have embarked upon. Vaismoradi et al. (2013) argue that the thematic approach is about the interpretation of data whereas content analysis focuses on quantifying data. This conclusion is arguably too simplistic, because Hsieh and Shannon (2005) present three approaches to qualitative content analysis. They started with ‘*conventional content analysis*’, here the primary aim was to describe a phenomenon inductively, without using any preconceived categories or ideas. The author contends that such a premise will be difficult to achieve because as soon as a researcher decides on a list of semi-structured questions there will be some constructivist bias. The next was the ‘*directed content analysis*’, where existing theory or prior research is used as the framework for understanding the phenomenon. Hsieh and Shannon (2005) support the findings of Mayring (2004) by positioning it as a deductive application to the qualitative process. This deductive and inductive differentiation was also presented by Braun and Clarke (2006) but their distinction was between the various forms of thematic analysis, although they concede that the inductive approach is similar to grounded theory. Hsieh and Shannon’s (2005) final option was the ‘*summative content analysis*’ approach, where text, images and other forms of data are quantified to explore usage as well as the interpretation of the information in question. Hsieh and Shannon (2005) did not explained if the ‘*summative content analysis*’ approach was inductive or deductive but the author contends that they could be either or a mixture of both.

Having considered the differences between the various content and thematic analysis options, the author will assert that the ‘*summative content analysis*’ or more specifically the ‘*passive, inductive, summative content analysis*’ approach is ideal for Netnography because the quantitative aspect of the analysis will give the researcher an idea of the dominant elements linked to the users/participants in question and the interpretivist aspect will provide a deeper/richer understanding of the problem. That said, researchers who form a pragmatic philosophical stance could support their findings with the statistical significance of their summative counts. Having completed this brief overview of the qualitative approach to

Netnography, the author will now demonstrate how an interpretivist approach to the ‘*passive, inductive, summative content analysis*’ can be applied to the domain in practice.

3. Methodology

As discussed earlier, the author has chosen to assess the perceptions of smart metres in the UK using a passive Netnographic approach (see Kozinets, 2020). The first task was to identify suitable sources of information to analyse. This search was limited to social media sites and internet forums. The author identified Mumsnet (www.mumsnet.com) as an ideal platform to initiate the analysis. “Mumsnet is now the UK’s biggest network for parents, with around 10 million unique visitors per month clocking up around 100 million-page views” (Mumsnet, nd). It had 88 separate threads [The author has defined a thread as a series of posts that are link to a question] that focused specifically on smart meters (i.e., where the initiating post had the words ‘smart meter(s)’ embedded within in it). The first ever post on smart meters was dated the 23/08/2011. Posts continued year on year, demonstrating an exponential profile with 37.5% (n=33 threads) occurring in 2019/20 (as at February 2020). The author made the decision to focus only on those messages published over the last 6 months. This was based on the argument that the most recent threads would give a better reflection on current perceptions of smart meters.

The author approached Mumsnet for permission to undertake a passive Netnographic review, which included the copying of a section of data onto NVivo. This was done to demonstrate good practice, even though the author would not be infringing copyright protection (see the ‘fair dealing’ exceptions identified in Sections 29 and 30 of the *Copyright, Designs and Patents Act 1988*). The author also confirmed that anonymity would be given to all participants. This meant that any direct quotes used in the final research report would need rewriting to ensure internet searches could not be traced back to a given participant. Ethics approval was also granted from the author’s research establishment. The coding of the data was completed in six phases, the overall framework can be seen in figure 1:

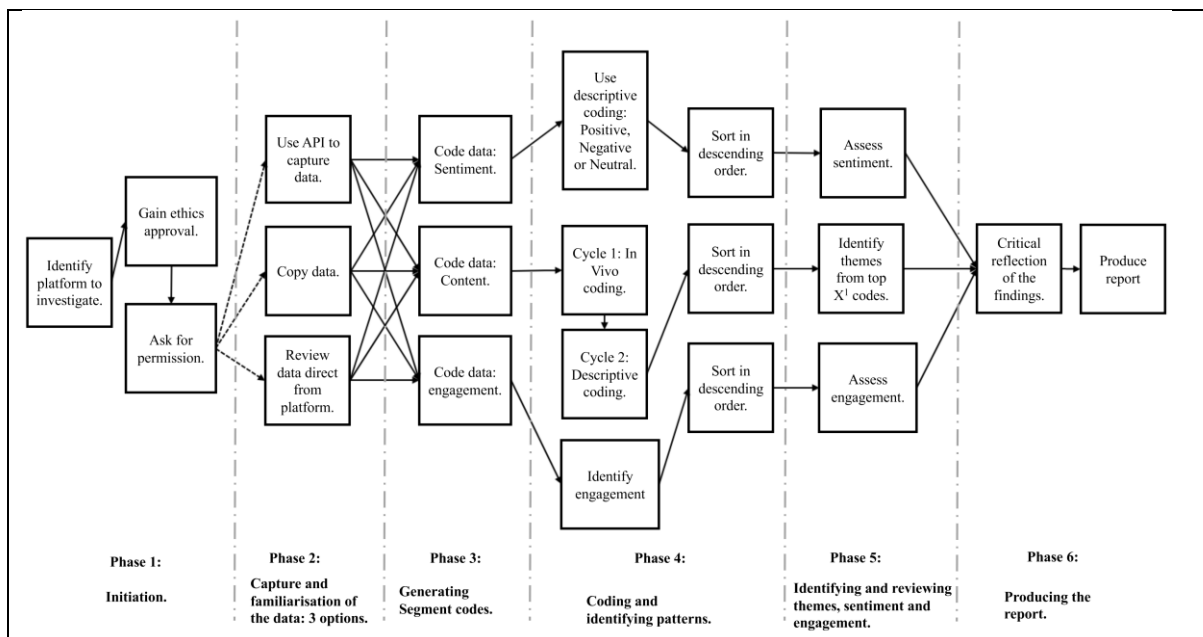


Figure 1: The passive Netnography framework using a summative content analysis approach.

Source: developed by Author. Note: X¹ is based on the Pareto Framework.

The first phase was the initiation and was discussed above. The second phase captured the data (note, the author only used the ‘copy data’ option), it also included the process of data familiarisation which is advocated by Braun and Clarke (2006) when embarking on any qualitative research. Phase three was the identification of segments to code, in this example three segments were identified. The first segment focused on sentiment (i.e., positive, neutral or negative). The second segment identified words, phrases and paragraphs for coding. The third segment identified the levels of engagement. Phase four was data coding and pattern identification. Here, the sentiment and engagement processes were done in one-cycle, but the theme development adopted a two-cycle approach: the author first used an ‘in vivo’ method to name codes (see Saldana, 2015). The second cycle compared all codes and their posts with the view of consolidating the data and creating a final code. This second cycle adopted a combination of Saldana’s (2015) ‘in vivo’ and ‘descriptive’ guidelines. As an example, after reviewing the initial codes of ‘wi-fi’, ‘broadband’ and ‘internet connection’ it had been established that the posts were arguably of the same family. As such they were given the final code of wi-fi. Where a code had no obvious relation with another, its final code remained the same as its initial code. The coding of sentiment and engagement used a ‘descriptive’ approach (see Saldana, 2015). For the sentiment analysis, each post was reviewed and was tagged as either positive or negative. The exception being the initial posts (i.e., those asking a question): they were classed as neutral. The engagement analysis focused on the thread profile, here the objective was to identify what type of thread gained the most engagement.

On completing the coding, the data was consolidated in NVivo and sorted into descending order, this was the start of phase 5. The counts for sentiment and engagement were taken, and the final theme development was initiated. Using Pareto’s 80:20 principle (see Sanders, 1987) only those codes within the top 80% of the cumulative post count were considered. Here an attempt was made to extend the final code, allowing a reader to have a better understanding of its context, as an example, the final code, ‘wi-fi’ was given the theme ‘needs a wi-fi signal’, which arguably provides the reader with a better understanding of what the researcher was intending to convey. Building on the work of Braun and Clarke (2006), the final themes and codes were analysed again to give the author an opportunity to select additional extract (i.e., those items not in the top 80%, but still have the ability to provide some insightful interpretations). The final phase was the critical reflection of the results and the production of the journal article.

4. Results and Discussion.

The messages reviewed consisted of 12 threads (i.e., the initial post/question) and 727 posts from 670 different participants, see table 1. The engagement (i.e., how many responses to an initiating question) had a mean and standard deviation of $M=60.58$ and $S.D.=77.22$ posts. The engagement values provide a market researcher with the opportunity to assess what type of question instilled the greatest response. This type of analysis is particularly useful if a researcher’s key question focused on what type of posts attract greater consumer involvement. In this instance, it can be seen that title’s with specific questions had a greater number of interactions, but there were more threads ($n=7$) with just the title ‘smart meters’.

Table 1: A summary of the post’s engagement.

Thread Title Code.	Threads	Posts	% of Posts	Users	% of Users
A question about people perceptions.	1	291	40.03%	278	41.49%
Smart meters.	7	206	28.34%	182	27.16%
A question about intimidation.	1	82	11.28%	73	10.90%
A question about smart meter problems.	1	73	10.04%	65	9.70%
Should I have a smart meter.	1	44	6.05%	42	6.27%
Inferring smart meters are now good	1	31	4.26%	30	4.48%

In terms of the sentiment analysis, a summary of the findings can be seen in table 2. To enhance the validity of the research, Sousa (2014, p.215) recommends that researchers “present interpretive evidence that should seek to go beyond citation” i.e., present examples of the posts so that readers can assess the thoughts of the researcher. The problem with adopting this approach in Netnography is the post will need altering (hopefully in a manner that does not distort the context). This is because it will mitigate the issues relating to anonymity, as the text cannot be used as a search term to identify the individual who made the post. It thus becomes a limitation of this research. Examples of the interpretations of the sentiment analysis are:

Negative sentiment: “I don’t trust the energy companies; I don’t see why I should have a smart meter.” (Participant 12)

Neutral sentiment: “Does anyone know anything about smart meters? Should I get one?” (Participant 1)

Positive sentiment: “I haven’t had a problem, they are helping me save money.” (Participant 57)

Table 2: Summary of the message sentiment.

Sentiment	Count	%
Negative	536	73.73%
Neutral	12	1.65%
Positive	179	24.62%
Total	727	

As mentioned in the methodology section, there is an argument that a researcher could adopt a statistical/quantitative approach in the analysis of the findings, the same could also be said for the differences associated with each theme and engagement. The author would argue that this would be perfectly valid, but the philosophical perspectives would need changing, a pragmatist stance would be the ideal philosophical paradigm for such an option. The author, however, has opted to remain an interpretivist and as such has not embarked on any statistical analysis. Earlier the author stated that a researcher should use the Pareto approach when summarising the themes, i.e., to present only those themes equating to 80% of the summative cumulative count. The final count can be seen in table 3, but it should be noted that researchers must be flexible in this process. Here, the author has expanded the boundary to just over 70% otherwise there would have only been 2-3 themes to discuss.

Table 3: Summary of the key themes identified.

Note: There are more than 390 themes because some messages had more than one theme attached to it.

Theme	Count	%	Cumulative %
Non-transferable	79	10.87%	10.87%
Not compulsory	49	6.74%	17.61%
Hacking concerns	40	5.50%	23.11%
Don’t bother	28	3.85%	26.96%
Needs a Wi-Fi Signal	17	2.34%	29.30%
Other Themes	514	70.70%	100.00%
Total	727		

Again, using Sousa (2014) guidance on increasing validity, the interpretation of the themes are as follows:

- a. **Non-transferable:** if participants changed their energy supplier at a later date, the smart meters would not work. A new one would have to be installed and it was unclear if they would have to pay for this. **Example:** “I had one, but when I changed supplier it wouldn’t work, what a joke.” (Participant 117).

- b. **Not compulsory:** a number of participants were under the impression that the fitting of smart meters was compulsory. This is not the case, households can decide if they want it installed. **Example:** *“If your provider says you must have one then they are lying, it is not compulsory, it’s up to you if you want to install one.”* (Participant 85)
- c. **Hacking concerns:** participants were concerned that the system would be hacked, and they would be charged for the wrong energy consumption. **Example:** *“I have heard that the systems can be easily hacked, I am not getting one until there is more assurance.”* (Participant 502).
- d. **Don’t bother:** there was a strong recommendation for participants to wait until the major issues identified are fully rectified before installing a smart meter. **Example:** *“The whole thing is a complete waste of time, it won’t save you any money and the contractors they use don’t know what they are doing: I had to stay at home on three separate occasions because the installers did not know what they were doing. I would wait until they get their act sorted.”* (Participant 304).
- e. **Needs a Wi-Fi Signal:** there was a misconception that smart meters require a wi-fi signal. **Example:** *“You need to have a wi-fi signal by the meter for it to work, ours is in the basement and it picks up nothing!”* (Participant 54).

The results demonstrate that there is a disconnect between consumers (those on Mumsnet at least) and the product, the UK Government should work with providers to make the product interchangeable between energy suppliers. The study identified that even the latest second-generation meters had problems with interchangeability. There also seems to be some misinformation about the product, 6.7% of the posts had assumed that the meters were compulsory and 2.34% believed that meters needed a wi-fi signal. These two misconceptions were wrong and could be rectified by better marketing. The hacking concerns and the conversations about ‘not to bother’ could also be addressed through better marketing.

The analysis provided above is only a brief overview of the findings, this is because the author’s main objective was to provide a critical evaluation of the issues relating to data capture and analysis in Netnography. The first of these issues can be linked to the Pareto model, the importance of themes outside of the designated threshold should not be wholly ignored. The use of Pareto’s 80:20 rule, would arguably enhance the summative content analysis process because it provides some rigour to the methodology and it addresses Hsieh and Shannon (2005, p.1285) concern that “the findings from this approach [summative content analysis] are limited by their inattention to the broader meanings present in the data”. Several other qualitative researchers advocate segmenting themes into major and minor groups (see Costa et al., 2017). An example is Houser (2019), although such a stance could be seen as cumbersome. The author thus advocates the consideration of minor themes in addition to those within the adapted Pareto frame, when applicable. It would ensure that a study is not overloaded with themes but has a range that provides a ‘rich understanding’ of the topic under review. It would also fit with the interpretivist and pragmatist paradigms so is unlikely to cause any philosophical conflict. These minor themes can still provide valuable insights into the perceptions of users. As an example, there was a theme tagged as “can save money”, which accounted for n=3 or 0.6% of the posts and represented an opinion that smart meters do work for some people. It demonstrates three points of interest, the first relates to the application of the model, in that there is a small cohort of users who have identified value in smart meters. The second relates to the sentiment analysis, it puts into context what people find attractive about the product. Finally, there is the question of influencer marketing, such individuals could be harnessed as advocates to support the brand/product or service, although the processes and procedures of such an approach go beyond the scope of this study. It does, however, demonstrate that Netnography can be used as a means

of identifying such individuals. This finding also demonstrates that adopting a prescribed ridged approach is not ideal in a qualitative Netnography process, which is a premise supported by Holloway and Todres (2003). That said, scholars must ensure that they can demonstrate inner consistency and coherence, which the author has interpreted as an ability of a piece of research to be replicated by others through a clear set of logical guidelines. This stipulation is arguably, similar to the guidance given by Sousa (2014) on the validation of qualitative research and has already been applied to this research.

The study identified 670 different participants had engaged in conversations relating to smart meters. The question that researchers should ask is, how many participants are required to validate the study? Crouch and McKenzie (2006) advocate 20 as an optimum number in qualitative research, but this is Netnography, where opinions are essentially snippets of information. So, can 20 posts be the minimum number to consider? To help resolve this dilemma, readers should consider the work of Sousa (2014) on the validation of qualitative research. He advocates the guidance from Morrow (2005), where a validity assessment can only be completed in relation to the research paradigm and epistemology. As mentioned in the methodology section, the author had adopted an interpretivist paradigm and a constructivist epistemology (see Gray, 2019). This means that results were based on the author's interpretation of the data using his experience as the key mechanism for analysis. As per Sousa (2014) guidance, readers should also consider the '*trustworthiness of the method*'. The question of trustworthiness increases (or should increase) when researchers present a detailed account of their methods, which the author has done. What is unique to Netnography, and is an element not considered by Sousa (2014), is the trustworthiness of the site under review. It could be that unscrupulous individuals may have signed up with multiple accounts and were set on misleading those reading the threads. To mitigate this, the study's author would need to view the IP addresses [An IP address (or Internet Protocol address) is a numerical label assigned to all devices connected to a computer network that uses the Internet Protocol for communication (Rooney, 2011)] of each participant, this was not possible, as such becomes another limitation of the study.

Conclusion

Taking all the above points into consideration, readers will hopefully see that an adapted passive summative content Netnographic approach can provide market researchers with insightful data about the topic in question. This is particularly important when access to the API is restricted. The author has also addressed the issues relating to copyright infringements and provided advice on how to mitigate against them. The passive aspect of embarking on this type of study would mean that there would be no engagement with users, making it a streamlined and cost-efficient way of assessing any product or service. It would also result in the ethics approval process becoming much easier. The adapted passive summative content Netnographic approach provides researchers with the ability to capture consumer thoughts and feelings without the risk of participant bias (see Goodwin and Goodwin, 2016) because there is no direct interaction with the researcher, although there are other limitations that researchers must consider like, being unable to confirm that each person in the study was unique (i.e., there were no unscrupulous bodies set on misleading others). Using the Pareto 80:20 rule would provide a researcher with the rigour needed to increase the validity of the research. That said, some flexibility is still needed to ensure the greatest number of appropriate insights are presented.

From the case study review, it was evident that there was a disconnect between consumers and the product. The UK Government and its smart meter providers should focus on how the functionality can be adapted such that it becomes interchangeable between different

energy suppliers. They should also work on their marketing campaigns to address the concerns identified which would also counter the misinformation that exists around the product.

Acknowledgments

The author would like to thank Mumsnet for providing the author with their permission to embark on this research and to use their name. Their forum can be found at www.mumsnet.com.

References

- BAND, J. and GERAFFI, J., 2013. The Fair Use/Fair Dealing Handbook. Available: <http://infojustice.org/wp-content/uploads/2013/03/band-and-gerafi-2013.pdf>.
- BARTL, M., KANNAN, V.K. and STOCKINGER, H., 2016. A review and analysis of literature on netnography research. *International Journal of Technology Marketing*, 11(2), pp. 165-196.
- BELZ, F. and BAUMBACH, W., 2010. Netnography as a method of lead user identification. *Creativity and Innovation Management*, 19(3), pp. 304-313.
- BRAUN, V. and CLARKE, V., 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), pp.77-101.
- CADWALLADR, C. and GRAHAM-HARRISON, E., 2018. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The guardian*, 17, pp. 22.
- COSTA, A.P., REIS, L.P., DE SOUZA, F.N. and MOREIRA, A., 2017. *Computer Supported Qualitative Research: Second International Symposium on Qualitative Research (ISQR 2017)*. London, UK, Springer.
- COSTELLO, L., MCDERMOTT, M. and WALLACE, R., 2017. Netnography: range of practices, misperceptions, and missed opportunities. *International Journal of Qualitative Methods*, 16(1), pp. 1-12
- CROUCH, M. and MCKENZIE, H., 2006. The logic of small samples in interview-based qualitative research. *Social science information*, 45(4), pp. 483-499.
- ELLIOTT, R., SHANKAR, A., LANGER, R. and BECKMAN, S.C., 2005. Sensitive research topics: netnography revisited. *Qualitative market research: An international journal*, 8(2), pp.189-203.
- GOODWIN, C.J. and GOODWIN, K.A., 2016. *Research in psychology methods and design*. Hoboken, NJ: John Wiley & Sons.
- GRAY, D.E., 2019. *Doing research in the business world*. London, UK: Sage Publications Limited.
- HARDY, S.A., 2017. Quantitative analysis of open-source data on metal detecting for cultural property: Estimation of the scale and intensity of metal detecting and the quantity of metal-detected cultural goods. *Cogent Social Sciences*, 3(1), pp. 1298397.
- HEINONEN, K. and MEDBERG, G., 2018. Netnography as a tool for understanding customers: Implications for service research and practice. *Journal of Services Marketing*, 32(6), pp. 657-679.
- HOLLOWAY, I. and TODRES, L., 2003. The status of method: flexibility, consistency and coherence. *Qualitative research*, 3(3), pp. 345-357.
- HOUSER, R.A., 2019. *Counseling and educational research: Evaluation and application*. London: Sage Publications.
- HSIEH, H.F. and SHANNON, S.E., 2005. Three approaches to qualitative content analysis. *Qualitative health research*, 15(9), pp. 1277-1288.

- KEIM-MALPASS, J., STEEVES, R.H. and KENNEDY, C., 2014. Internet ethnography: A review of methodological considerations for studying online illness blogs. *International journal of nursing studies*, 51(12), pp. 1686-1692.
- KOZINETS, R.V., 1998. On netnography: Initial reflections on consumer research investigations of cyberculture. In: J.W. ALBA and J.W. HUTCHINSON, eds, *Advances in Consumer Research Volume 25*. Provo, UT: Assoc for Consumer Research, pp. 366-371.
- KOZINETS, R.V., 2010. *Netnography: Doing ethnographic research online*. Thousand Oaks, CA: Sage Publications Limited.
- KOZINETS, R.V., 2015. *Netnography: Redefined, 2nd ed.*, Thousand Oaks, CA: Sage Publications Limited.
- KOZINETS, R.V., 2020. *Netnography: The Essential Guide to Qualitative Social Media Research*. London, UK: Sage Publications Limited.
- KOZINETS, R.V., DOLBEC, P. and EARLEY, A., 2014. Netnographic analysis: Understanding culture through social media data. In: U. FLICK, ed, *The SAGE Handbook of Qualitative Data Analysis*. London, UK: Sage Publications Limited. pp. 262-276.
- KROTOV, V. and SILVA, L., 2018. Legality and ethics of web scraping, *Twenty-fourth Americas Conference on Information Systems*, 2018.
- LAURET, A., 2019. *The Design of Web APIs*. Shelter Island, NY: Manning Publications.
- LEXISNEXIS, nd. Obligations to transpose and process EU directives. Available: <https://www.lexisnexis.co.uk/legal/guidance/obligations-to-transpose-and-process-eu-directives>.
- LUGOSI, P., JANTA, H. and WATSON, P., 2012. Investigative management and consumer research on the internet. *International Journal of Contemporary Hospitality Management*, 24(6), pp. 838-854.
- MAYRING, P., 2004. Qualitative content analysis. In: U. FLICK, E. VON KARDOFF and I. STEINKE, eds, *A Companion to Qualitative Research*. London: Sage Publications Limited. pp. 266-269.
- MORROW, S.L., 2005. Quality and trustworthiness in qualitative research in counseling psychology. *Journal of counseling psychology*, 52(2), pp. 250.
- MRS, 2014. MRS Guidelines for Online Research. Available: <https://www.mrs.org.uk/pdf/2014-09-01%20Online%20Research%20Guidelines.pdf>.
- MUMSNET, nd. About Us. Available: <https://www.mumsnet.com/info/about-us>.
- OFGEM, 2019. *Statutory consultation on the post-2020 smart meter rollout supplier reporting requirements*. Available: https://www.ofgem.gov.uk/system/files/docs/2019/10/statutory_consultation_on_the_post-2020_smart_meter_rollout_reporting_requirements_0.pdf.
- PUSCHMANN, C. and AUSSERHOFER, J., 2017. Social Data APIs: Origin, Types, Issues. In: M.T. SCHÄFER and K. VAN ES, eds, *The datafied society studying culture through data*. Amsterdam, Netherlands: Amsterdam University Press, pp. 147-154.
- REID, E. and DUFFY, K., 2018. A netnographic sensibility: Developing the netnographic/social listening boundaries. *Journal of Marketing Management*, 34(3-4), pp. 263-286.
- ROONEY, T., 2011. *IP Address Management: Principles and Practice*. Hoboken, NJ: John Wiley & Sons.
- SALDAÑA, J., 2015. *The coding manual for qualitative researchers*. Third ed. London: Sage Publications Ltd.
- SANDERS, R., 1987. The Pareto principle: its use and abuse. *Journal of Services Marketing*, 1(2), pp. 37-40.

- SOUSA, D., 2014. Validation in qualitative research: General aspects and specificities of the descriptive phenomenological method. *Qualitative Research in Psychology*, 11(2), pp. 211-227.
- VAISMORADI, M., TURUNEN, H. and BONDAS, T., 2013. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & health sciences*, 15(3), pp. 398-405.
- VENTURINI, T. and ROGERS, R., 2019. “API-Based Research” or How can Digital Sociology and Journalism Studies Learn from the Facebook and Cambridge Analytica Data Breach. *Digital Journalism*, 7(4), pp. 532-540.