Four styles of quali-quantitative analysis

Making sense of the new nordic food movement on the web

Munk, Anders Kristian

Published in:
Nordicom Review

DOI (link to publication from Publisher):
10.2478/nor-2019-0020

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Four Styles of Quali-Quantitative Analysis

Making sense of the new Nordic food movement on the web

Anders Kristian Munk

Department of Learning and Philosophy, Aalborg University, Denmark, akm@learning.aau.dk

Abstract

Through the example of a web corpus built to study the emergence of the New Nordic Food phenomenon in Scandinavia, I discuss how quali-quantitative analysis can help us make sense of onlife traces. I propose four styles of analysis that address the meaning problem in different ways, namely 1) through *complementarity*, a division of labour in which quantitative and qualitative methods are allowed to unfold relatively undisturbed by one another, the latter performing the job of situating and interpreting the insights gleaned from the former; 2) through a *single level of analysis*, whereby the potential of onlife traces is seen to reside in their ability to be both qualitatively rich and quantifiable at the same time, enabling an analysis of how apparent macro phenomena are produced on the micro level; 3) through *curation*, a critical practice in which a qualitative understanding of different media environments and their effects on the production of onlife traces becomes integral to the way in which such data should be sourced and quantified; and 4) through *algorithmic sensemaking*, whereby the relational reasoning typically associated with qualitative fieldwork is emulated quantitatively through techniques like pattern recognition.

Keywords: digital methods, mixed methods, ethnography, big data, sensemaking

Introduction

> I came back duly, and soon gathered an audience around me. A few compliments in pidgin-English, some tobacco changing hands, induced an atmosphere of mutual amiability. I tried then to proceed to business. First, to begin with subjects which might arouse no suspicion, I started to "do" technology. A few natives were engaged in manufacturing some object or other. It was easy to look at it and obtain the names of the tools, and even some technical expressions about the proceedings, but there the matter ended. (...) I took a village census, wrote down genealogies, drew up plans, and collected the terms of kinship. But all this remained dead

Munk, A. K. (2019). Four styles of quali-quantitative analysis: Making sense of the new Nordic food movement on the web. *Nordicom Review,* 40 (Special Issue 1): 159-176. doi:10.2478/nor-2019-0020.

> material, which led no further into the understanding of real native mentality or behaviour, since I could neither procure a good native interpretation of any of these items, nor get what could be called the hang of tribal life. (Malinowski, 1922: 5-6)

In the opening paragraphs of *Argonauts of the Western Pacific*, Bronislaw Malinowski (1922) set out a core methodological problem that has occupied ethnographers in the field for more than a century. Whereas it is relatively straightforward, through conversation and direct observation, to record everyday life in a social setting, it is much harder to obtain what Malinowski (1922: 24) called a "native's point of view". Having the data is, of course, never the same as understanding what they mean, but ethnographers must also understand what they mean to their informants. They routinely reckon with the fact that knowledge of how humans ascribe meaning to their everyday lives does not come simply from making sense of data after they have been collected but from ensuring, as part of the empirical work in the field, that sensemaking happens in accordance with native theories of the world. It is thus both a systematic account of a social setting and an interpretation of the situations described in this account that must be obtained from the group of people under study.

This version of the meaning problem is arguably peculiar to ethnography. Indeed, ethnographers have developed their own vocabulary for it, such as the distinction between *emic* and *etic* knowledge (i.e. native and anthropological theories of the world) or the notion of *ethnocentrism*. However, it is, at the very least, a version of the meaning problem that has strong affinities with the situation in digital sociology (as defined by e.g. Marres, 2017; Rogers, 2013). When we try to make sense of onlife traces, there are always unresolved questions about how these traces were intended, what they meant in their original context and, indeed, what that context was. These questions are not simply about making sense of data but about making sense in a way that claims to reflect a native point of view. Like Malinowski's dutiful but dissatisfying first attempts at charting Trobriand culture, we too can curate social media collections, record metadata and map digital interactions without coming any closer to getting the "hang of tribal life" (Malinowski, 1922: 6).

To illustrate the nature of this problem, I draw on the example of a web corpus (a collection of websites connected by their hyperlinks) that I curated in 2012 as part of a project that aimed to map the so-called New Nordic Food movement and trace its development across the Scandinavian countries. The New Nordic Food movement formally traces its origins to the publication of the *Manifesto for the new Nordic kitchen* in 2004. In this manifesto, a group of chefs and restaurateurs from Denmark, Sweden, Norway, Finland, Iceland, Greenland and the Faroe Islands signed up to 10 future goals.[1] Rather than inventing a phenomenon from scratch, the signatories of the manifesto were playing to a range of established agendas, each with their own histories and trajectories, which were, at the time, relatively unrelated to one another. Besides the gastronomic quest to create a local and seasonal Scandinavian cuisine that would be able to rival its counterparts on the global culinary scene (in itself a considerable challenge), the Manifesto expressed ambitions in areas like public health, regional development, sustainability, animal welfare, Nordic identity building and cooperation, local self-sufficiency and innovation in the food and tourism sectors. However, it was not at all obvious how this new Nordic issue amalgam could be forged practically. Despite considerable, long-

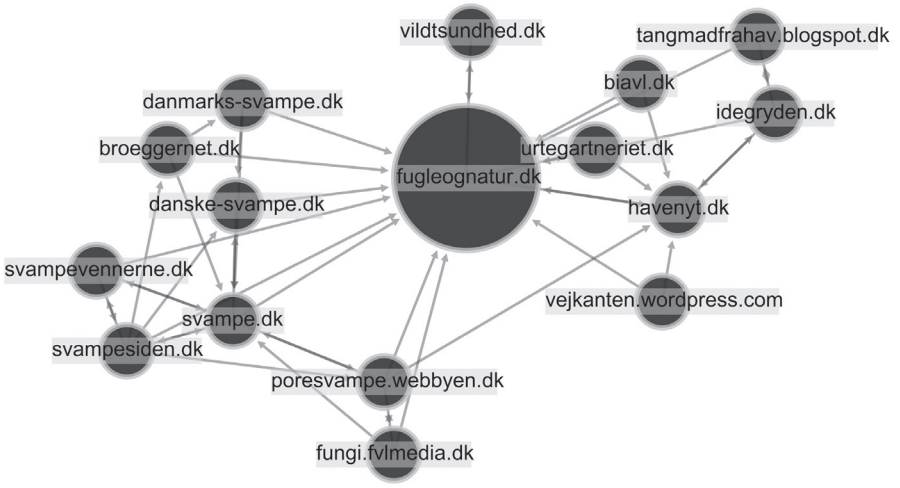
standing grassroot efforts around organic food production and sustainable development, the Manifesto came at a time when it was still uncommon to find vegetables, fish, meat or dairy products of any local typicality in ordinary retail in Scandinavia; terroir was something you might look for in an imported bottle of French wine but certainly not in your carrots or your cream cheese. Much less were public health narratives attuned to the virtues of a locally sourced diet – the Mediterranean "miracle" with its exotic whiff of southern *art de vivre* figured more as an interesting than as a potential model to follow – and rural development was typically not associated with top-tier cherry wine, hand-picked sea urchins or the recultivation of ancient cereals like emmer or einkorn wheat. The year before the publication of the Manifesto, Claus Meyer and René Redzepi, two of the signatories, had opened Noma in Copenhagen with the stated ambition of basing a restaurant exclusively on seasonal ingredients from the Nordic region. The fact that foie gras and olive oil would be banned from the menu and that one would instead be able to feast on musk ox, reindeer, malted barley or butter milk was met with widespread scepticism in culinary circles, which were quick to nickname the restaurant "the seal fucker" ("sælpuleren" in Danish) in a sarcastic gesture towards its North Atlantic pantry.

Eight years later, when I was building my web corpus, that situation had dramatically changed. Following up on the Manifesto, the Nordic Council of Ministers had launched its New Nordic Food programme to promote awareness of the Nordic region as a gastronomic destination and encourage cross-sectoral innovation and business development. In 2009, the Faculty of Life Sciences at the University of Copenhagen had received a research grant of €13.3 m for a project on "optimal well-being, development and health for Danish children through a New Nordic Diet" (OPUS). In addition, in 2010, Noma had dethroned El Bulli in Catalunya as the World's Best Restaurant on the annual top fifty list published by the British Restaurant Magazine, a feat that it repeated three years in a row, bringing a veritable surge of young chefs and fine-dining lovers to Copenhagen in search of the new Nordic revolution. Similar restaurants had opened up in Sweden and Norway and were climbing quickly up the charts of the international gastocracy. Importantly, it was not only the chefs but also their growers and suppliers that attained star status. Lifestyle magazines were writing stories about a former investment banker from London who was handpicking sea urchins in Arctic Norway, an eccentric asparagus grower by a fjord on Zealand and a wildly bearded forager from Scania. Such stories, and many others, seemed to cement the impression that something had changed fundamentally in the way in which Scandinavia perceived itself (and was being perceived by the rest of the world) as a gastronomic hot spot.

From an ethnographic point of view, what made the New Nordic Food movement interesting and worthy of inquiry was the obvious differences in the way in which it had been translated into everyday practice across Scandinavia. Even from a position in Copenhagen, where it was common to be presented with the same rather heroic epic about Nordic chefs conquering the world, there were differences to be observed, for example between the intervention research taking place at the Faculty of Life Sciences and the fermentation experiments happening in the kitchen at Noma (see also Munk & Ellern, 2015). Nevertheless, high gastronomy and public health had somehow become related. It was clear that changes were happening to the social fabric and that these changes were likely to take on very different forms across Scandinavia. Figure 1 shows the beginning of one of my attempts to map these differences. I used the Navicrawler plug-in for

Firefox to collect websites that are in one way or another dedicated to foraging, wild foods and gastronomy in the Nordic region. The figure shows the resulting network of connections between Danish mushroom and wild herb enthusiasts.

Figure 1. *The beginning of a web corpus forming around the website of www.fugleognatur.dk (ego network)*



Comment: The websites were crawled using the Navicrawler plug-in for Firefox (Jacomy et al., 2007) and they were spatialized in Gephi with ForceAtlas2 (Jacomy et al., 2014) and sized by their in-degree (the number of hyperlinks received from other sites in the network).

The Navicrawler browses through the pages of each website and indexes the hyperlinks. From this neighbourhood of linked websites, the researcher can make choices about where to visit next. The plug-in is specifically designed to keep the crawling process on a tight manual leash (Figure 4 gives a sense of the interface). Every time a site is crawled, the user must decide which of the neighbouring sites discovered through outlinks to include in the corpus and in turn crawl. I gradually developed the idea that I would continue this iterative process as long as I was discovering new food-related websites from Scandinavia. Eventually, after months of work, I ended up with a data set of 2007 websites that I had manually vetted as being both on the topic and located in the region and from which I had scraped all the hyperlinks (as shown in Figure 6) and was able to search for the presence of particular indicator words (as shown in Figure 2).

This collection of hyperlinks and indicator words of course consists of onlife traces. They tell us something about who these actors are associating with (from Figure 1 we can already begin to discern two distinct communities), what they are talking about and when they do so (as shown in Figure 2). However, they do not reveal the meaning of those actions to us. They constitute something akin to what Malinowski began drawing up when he first arrived in the Trobriands, namely the readily observable stuff. They present us with a similar problem to that facing Malinowski, namely how to achieve a native interpretation of that stuff. The barriers that prevent us from achieving this present themselves in different ways. For the ethnographer committed to long-term fieldwork in an unfamiliar setting, they are typically manifested as a frustrating inability to become

accepted by the group and understand the cultural codes that govern social interaction. For the digital sociologist, they are felt rather as a lack of context. Engaging with onlife traces involves an unsettling absence from the social situation of the informants.

Conventionally, in ethnography at least, it is the presence in the field that affords the researcher an opportunity to gauge whether he or she is making sense of the world on the same terms as the informants. That affordance is somewhat lost when working with onlife traces. We do not feel the inability to make sense directly as a problem in the same way as, for instance, Clifford Geertz described it when, together with his wife, he tried to establish rapport with their Balinese informants: "As we wandered around, uncertain, wistful, eager to please, people seemed to look right through us with a gaze focussed several yards behind us on some more actual stone or tree" (Geertz, 2005: 56). The fact that researchers working with onlife traces are not, in this way, faced with the consequences of not making sense arguably compounds the meaning problem. This is because there are no immediate social repercussions to remind the researcher that the problem exists. One could easily (in fact is highly likely to) gaze at the tightly connected cluster of Danish mushroom pickers on the left of Figure 1 and imbue it with meaning, for example that mushroom pickers respect each other's advice and thus associate intensely, without checking such an interpretation with the mushroom pickers themselves.

There is a striking similarity, however, in the disconnect that both ethnographers and digital sociologists experience between the ease with which they can chart and observe traces of social life and the difficulties involved in achieving a native perspective on whatever has been charted and observed. Indeed, the similarity extends to the ways in which we attempt to overcome those difficulties in our analyses. By taking us back to Malinowski, I deliberately gesture towards an inflection point in the disciplinary history of social anthropology at which in-depth, qualitative work came to be seen as the solution to the meaning problem. Ethnographers like Malinowski were brought up as part of a larger and far more quantitative collection effort organized, among others, by the Cambridge school of anthropologists under the leadership of Alfred Cort Haddon, Charles Seligman and William Halse Rivers. In the late nineteenth and early twentieth centuries, considerable scientific effort was devoted to the systematic collection of ethnographic field data, for example through standardized anthropological questionnaires (Munk & Jensen, 2015; Urry, 1972). We could thus say that Malinowski embodied the moment at which quantitative survey methods and qualitative participant observation intersect in anthropology. Digital methods are arguably in the middle of a similar moment, when a combination of qualitative and quantitative work is broadly construed as the solution to the meaning problem. However, whereas in social anthropology this moment came to signify an almost all-encompassing shift from surveys and questionnaires to in-depth interviews and participant observation, the study of onlife phenomena seems to offer new and interesting opportunities for fusing and merging the qualitative and the quantitative. What makes onlife phenomena fascinating in this respect is their potential for being both data traces and venues for qualitative work at the same time. One of the widely cited promises of digital sociology is thus the capacity to bridge the quali-quantitative divide (e.g. Latour et al., 2012).

My purpose in this article is to explicate some of the rather different ways in which that might happen. We know from mixed-methods research that the marriage of qualita-

tive and quantitative methods raises questions about how we collect data, what we do with those data once collected and how we can think more philosophically about those processes (Johnson et al., 2007; Tashakkori & Creswell, 2007). Is the quali-quantitative potential to be found in the fact that onlife traces are often both quantifiable (counting shares and likes or developing network metrics for interaction) and qualitatively rich (comprising unstructured text, user tags, profile data and the ability to situate within online communities)? Is it to be found in the way in which we engage with code and numerical data in more ethnographic or exploratory ways? Does it lie in our newfound capacity to make datascapes (i.e. interfaces that make messy and unstructured data navigable) available for exploration by our readers or our informants (Girard et al., 2016; Venturini et al., 2017)? Alternatively, is it to be found in the questions that we ask about the media environments that constitute "the field"?

Drawing on my example of the New Nordic web corpus, I identify four styles of quali-quantitative analysis that could all be said to address the meaning problem in their own specific way. The first, which I dub *complementarity* (following Bryman, 2006; Greene et al., 1989), adopts a position quite similar to what Huey Chen (2006: 75) called "program theory" in mixed-methods research. It assumes a basic division of labour in which the quantitative mapping of websites and hyperlinks through crawling and subsequent network analysis is allowed to produce insights on its own terms, which then complement qualitative interviews in the field that are also allowed to unfold on their own methodological premises. The hyphen between quali and quanti in this style of analysis simply signifies the completion of the picture. The two approaches contribute valuable insights but must be left to their own devices to do so, the premise being that the world views of qualitative and qualitative researchers are so fundamentally different that they should be allowed to co-exist undisturbed. It is thus the task of the project manager to ensure the right conditions (establish a "program") for this to happen.

Table 1. *Four styles of quali-quantitative analysis*

The quali-quantitative as:	Addresses the meaning problem by:
Complementarity	Interpreting insights from a quantitative analysis of onlife traces by situating these insights qualitatively in the everyday environments that they claim to reflect.
Single-level analysis	Tracing how quantitative patterns on the macro level emerge directly from qualitatively rich interactions on the micro level, demonstrating how onlife traces simultaneously embody both qualitative richness and quantifiability.
Curation	Critically reappropriating (and thus manually curating) onlife traces to speak on behalf of certain phenomena or address certain questions.
Algorithmic sensemaking	Attempting to reveal emic modes of ordering the world through quantitative techniques, for example pattern recognition.

The second style, which I dub *single-level analysis*, draws on the work of Michel Callon and Bruno Latour in actor–network theory (ANT) to overcome the distinction between micro and macro levels of analysis (Callon & Latour, 1981). What becomes possible on a website is precisely to trace quantifiable items like words or hyperlinks

directly back to the qualitative contexts in which they were deposited. This allows an ANT-inspired analysis to realize a kind of inquiry in which macro phenomena are explained by showing how they are constructed on the micro level. The third style, *curation*, takes seriously an idea proposed by Richard Rogers in his foundational work on digital methods, namely that we, as onlife researchers, must develop our own "critical analytics" (Rogers, 2018) that are tailored to our research questions. The native "vanity metrics" of the Web must thus be "repurposed", and this requires a qualitative intervention.

Finally, as a matter of discussion, I propose a fourth style of analysis that I call *algorithmic sensemaking*. It is fundamentally different from the other styles in that it does not involve any conventionally qualitative work but rather solicits sensemaking for quantitative community detection and pattern recognition. Here the hyphen between quali and quanti signifies a specific denomination of quantitative methods that exhibit some of the exploratory affordances otherwise attributed to qualitative work.

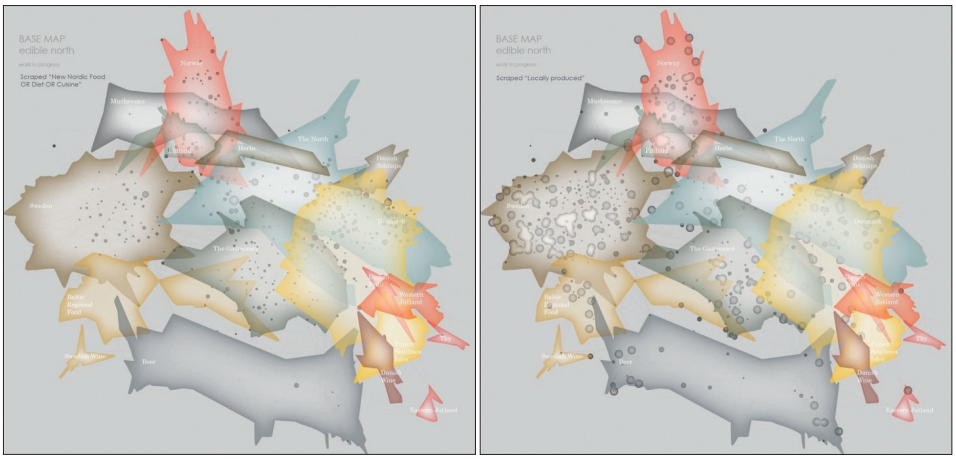
The quali-quantitative as complementarity

The most obvious way to understand the hyphen in the quali-quantitative is to let it signify a complementary relationship (Blok & Pedersen, 2014). If that is the case, then the quantitative and the qualitative can be allowed to play their separate roles, performing work that is already familiar to us, and leaving it to the researchers to argue from situation to situation how they can enrich each other. The gist of such an argument is often that in-depth ethnographic sensemaking is necessary to complete the picture in a world where machine learning and social physics have become ever more omnipresent. The most straightforward way to approach the quali-quantitative as a question of complementarity is thus through a division of labour. As Tricia Wang put it in a now much-cited blog post:

> Big Data reveals insights with a particular range of data points, while Thick Data reveals the social context of and connections between data points. Big Data delivers numbers; thick data delivers stories. Big data relies on machine learning; thick data relies on human learning. (Wang, 2013)

To demonstrate how big and thick data can work in mutual support of each other, I take my mapping of the New Nordic Food movement as an example. One of the first explorations of the web corpus consisted of a query. I simply searched the data set for mentions of terms relating to New Nordic Food. In this way, it was possible for me to show not only how the usage of the concept had migrated into different contexts but also how it compared to tangential concepts, like organic, biodynamic, locally produced, slow food or terroir. Figure 2 shows a graphic elaboration of the full corpus of 2007 websites clustered into different communities. The clusters in the network have been replaced by coloured polygons, and a semantic heatmap is overlaid on top. The heatmap shows where a given query term is found and how often (Figure 2 compares "new Nordic" with "locally produced"). It thus visualizes the results of a quantitative analysis (clustering by modularity and force vector layout, heatmaps by word count) of onlife traces (hyperlinks and text).

Figure 2. *Stylized interpretation of the web corpus (2007 websites connected by hyperlinks – see also Figure 6) with two semantic overlays*



Comment: On the left, mentions of the terms "new Nordic food", "new Nordic cuisine" or "new Nordic diet" in any of the Nordic languages are highlighted. On the right are mentions of the term "locally produced".

This quantitative analysis is interesting on its own merits, because it prompts questions. Notice for example that the Swedish cluster (coloured brown on the left side of both maps) not only has far more talk about local production (the map on the right) than anything new Nordic (map on the left) but that the left half of the cluster makes no mention of new Nordic at all. Websites in this part of the cluster are typically producers, local authorities and other local food actors; websites in the inner parts of the cluster, where the New Nordic Movement is cited, are typically governmental or otherwise national in scope. In the Norwegian cluster (coloured red at the top of both maps), the pattern is similar (there is far more talk about local production), although here the New Nordic Food movement is cited specifically by a group of actors in the region of Trønderlagen.

A clear pattern thus emerges when we map onlife traces at scale, and it seems to be telling a story about the emergence of a new Nordic agenda and about the pre-existence of various other tangential food agendas, but we cannot discern the plot of this story or begin to understand its characters without further work. We can of course continue to refine the mapping. There is still a range of more quantitative, data-related questions to pursue. We need to consider how to normalize the raw occurrence of a query term relative to the size of the website on which it is found, for instance, and we could triangulate the findings, through other online media and/or other query methods, to corroborate that there is indeed a pattern. Being able to do so, however, will not provide an interpretation of why the pattern is there.

This is where complementary qualitative work becomes useful. Prompted by the questions raised by the maps in Figure 2, I decided to visit some of the regional food actors in Trondheim. In the offices where my first interviews took place, I noticed a framed version of the *New Nordic food manifesto* hanging on the wall. This version had a slightly different title: it was called the *Trønderlagen food manifesto* (see Picture 1), but the bullet points were more or less identical to the original *New Nordic food manifesto*. Aslaug, who I was due to meet, explained its provenance. In 2009, her agency, which

was tasked with promoting food products from the Trønderlagen region, had invited Claus Meyer, one of the instigators of the original manifesto, to act as a consultant on how to develop and grow the local food sector. The agency had since adopted the principles of the Manifesto as a template for achieving this goal. "The idea is not to have something to simply hang on the wall", Aslaug told me, "we also counsel schools and municipalities on how to use more local products".

Picture 1. *Trønderlagen's food manifesto in the offices of a food agency in Trondheim*

Later that day, while visiting a monthly food fair in Trondheim, I talked to a range of local producers and experienced a very different story. Be they dairy farmers, reindeer herders or the woman selling smoked whale meat from the fjords outside the city, none of them were the least familiar with the concept of New Nordic Food. Rather, they talked about their products as small scale and locally sourced. It reminded me of something Aslaug had told me earlier that day. Small-scale local production has always been a key part of the Norwegian agricultural policy. From Kristiansand in the South to Krikenæs in the North, the Norwegian Government works continuously to ensure that food can be produced locally, even if it would be cheaper to import it from abroad. This is a policy that is about national and local self-sufficiency in a country that stretches more than 25,000 km, all the way north of the Arctic Circle. It has kept local production at scales that would have been unimaginable in Denmark, where most small-scale producers were merged into larger and larger cooperatives in the latter half of the twentieth century. Trondheim's local production has a completely different history and a different set of

connotations from that in Copenhagen. Whereas the New Nordic Food movement in Denmark has come to signify a transition back to local food production, in Norway it is really more a question of rebranding small-scale producers and encouraging them to grow into viable export businesses that would help to develop the economy. New Nordic is seen as a development policy rather than as a culinary resuscitation project. This provides an explanation for why the concept does not (yet) occur in the discourse of the Norwegian producers on the maps in Figure 2.

To think about the quali-quantitative as a complementary relationship thus reflects, in the case above, a way of interpreting the findings from a quantitative analysis of onlife traces through additional qualitative work in the field. Although much mixed-methods research has generally adhered to this idea of existing methods complementing each other's strengths and weaknesses, it is worth noting that the sequence in which it happens is typically reversed. In what Creswell called "exploratory sequential" research designs (Cresswell, 2014: 6), it is an initial qualitative analysis that produces questions and hypotheses, which can then be tested and affirmed through, for instance, a quantitative survey. The meaning problem, as it arises specifically in relation to the study of onlife traces and historically in much ethnographic work, tends to (or at least can very easily) reverse that sequence. The issue is not (only) how qualitative findings can be made more robust through additional quantitative work but how quantifiable signals were intended and how emic sense can be made of them.

The quali-quantitative as a single level of analysis

When the data set is a corpus of websites, opportunities also arise for performing qualitative work directly on that data set. Even though the complementary division of labour that I outlined above is clearer when the qualitative work takes place in the field and the quantitative operations take place on data, the fact is that many onlife traces are both deep and big at the same time. A website can simultaneously be a place to conduct discourse or image analysis, count words or hyperlinks and, if it hosts a forum, a place to hang out and interact with users. This could be leveraged simply as another way to let qualitative and quantitative approaches complement each other. Qualitative work on websites provides depth to patterns detected in data from these sites. Many virtual anthropologists thus make use of automatic data capture tools to support their otherwise qualitative exploits online (Boellstorff et al., 2012).

Arguably, however, the simultaneous depth and breadth that characterize onlife traces also offer an opportunity to overcome the quali-quantitative divide. In making this argument, Bruno Latour (2012) echoed a decade-old actor–network theoretical dream of an analysis that does not a priori separate the world into a micro and a macro level (compare Callon & Latour, 1981 with Latour et al., 2012). He argued that onlife traces are rapidly making such a dream possible:

> The best proof that those two levels do not correspond to any real ontological domains is that they begin to disappear, to be literally redistributed, every time one modifies or enhances the quality of access to the datasets, thereby allowing the observer to define any actor by its network and vice versa. This is exactly what the striking extension of digital tools is doing to the very notions of "individual"

> and "wholes". The experience (more and more common nowadays) of navigating on a screen from elements to aggregates may lead researchers to grant less importance to those two provisional end points. Instead of having to choose and thus to jump from individuals to wholes, from micro to macro, you occupy all sorts of other positions, constantly rearranging the way profiles are interconnected and overlapping. (Latour et al., 2012: 595)

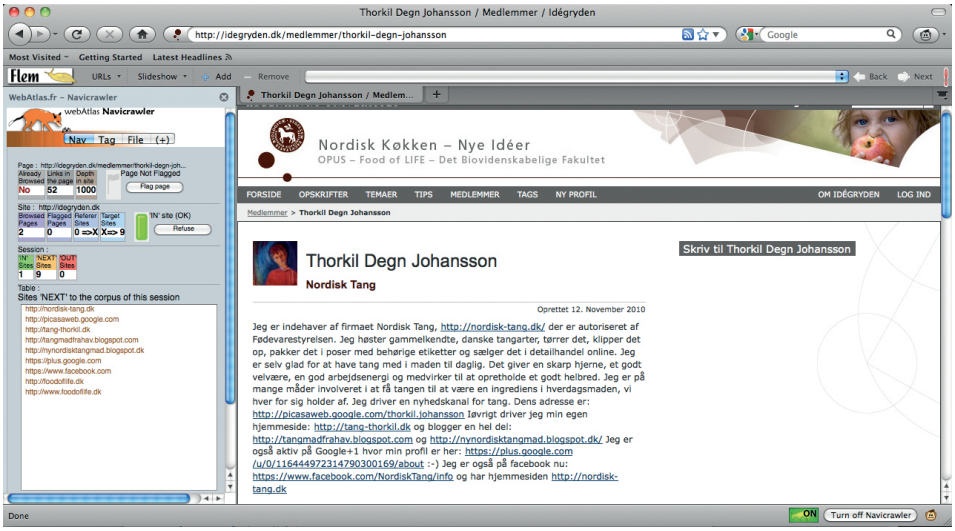
Data-intensive methods are the means to an end here, not a separate mode of analysis. They make it possible to aggregate and disaggregate interactions without introducing a simplified macro level (e.g. "the social", Latour, 2005) that bears no direct empirical connection to these interactions (which are then confined to their own micro level of analysis). It was Gabriel Tarde (2015) who first proposed the idea of a so-called "monadic" sociology. As opposed to Émile Durkheim (who eventually supplanted him), Tarde refused to resort to abstract referrals to social laws that must be assumed to exist *sui generis* in the service of a social organism. Instead, he envisioned that a corps of ethnologists would travel around France and systematically record the details of everyday life until emergent patterns of imitation could be discerned (Barry, 2010). The tools to manage that sort of pattern extraction were not available in the late nineteenth century, and those details of everyday life had not been datafied to any significant extent. This can be said to have changed with online media and computational methods. What is becoming possible is a version of the quali-quantitative that scales smoothly from minute details to aggregate patterns and back.

The difference between a complementary approach and a single level of analysis approach can again be illustrated through the example of the New Nordic web corpus. One of the interesting features of this corpus is a particular group of tightly connected websites that display an intense interest in seaweed. A complementary approach could notice this feature on the macro level through quantitative techniques and then attempt to make sense of it on the micro level through in-depth qualitative work. Figure 4 shows a screenshot from one of the websites in question. It allows us to situate the seaweed talk in context, in this case on the profile page of a participant in the OPUS research project who nicknames himself "Seaweed Thorkil". We can for example begin to build an interpretation that involves arguments about the health benefits of seaweed (which makes it particularly interesting to actors carrying out research on the New Nordic Diet) and its value as a sustainable food source.

For a single-level analysis, however, the profile page of Seaweed Thorkil is not so much an opportunity to interpret the meaning of a pattern as it is an opportunity to show how this pattern is wholly contingent on work that happens at the micro level (thus blurring the micro/macro distinction). The pattern that we detect at scale can be traced in part to how Thorkil shares hyperlinks, how he writes about seaweed and how the websites of projects like OPUS make these sharing practices possible.

The American anthropologist George Marcus (1995) reminded us that the ethnographer is not only in the business of providing thick descriptions but should also take responsibility for positing the specific logics of association that define the space of research (the context within which cultural analysis takes place and makes sense of the world). For too long, he argued, this task of delimiting the field has been allocated to third parties in the academy, for instance historians, geographers or economists, conveniently allow-

Picture 2. *The web corpus under construction (Navicrawler is running as a Firefox plug-in with the control panel to the left)*



Comment: Screenshot from the website of www.idégryden.dk, the online forum of the OPUS research project exploring the potentials of the New Nordic Diet.

ing ethnography to maintain its self-image as the branch of social science that provides depth and meaning rather than overviews (maintaining the quali-quantitative divide, we could add). Marcus was drawing on, and tuning into, a growing body of anthropological scholarship that was questioning the isomorphic conflations of "place", "culture" and "field" that would more or less automatically prescribe a geographically bounded space of research, even in a world where flows and interconnectivity had become the order of the day (e.g. Appadurai, 1988; Olwig & Hastrup, 1997). Without attempting to cover the wide range of interconnective circumstances under which this isomorphism was found to be failing, it suffices to say that the advent of electronic mediation, and the internet in particular, was definitely among them.

Marcus foresaw that a set of methodological anxieties would be likely to result from anthropologists abandoning their spatially confined, physically locatable, single-site notions of the field. They were, in order of appearance, 1) that the limits of ethnography would be tested, because the attempt to grasp some wider system or context than a specific cultural locale situated within it might turn out to be better served by methods imported from other domains of the social sciences; 2) that the power of fieldwork would be attenuated, because the intensity of the ethnographic experience and the commitment to classic ethnographic virtues would suffer from the researcher's newfound mobility and ensuing lack of embeddedness; and 3) that the subaltern would be lost, that is, that the traditionally privileged position in ethnographic accounts of those placed outside the hegemonic power structures of the world system would gradually be eroded by attempts to "study up" or otherwise "complete the picture" (ibid.).

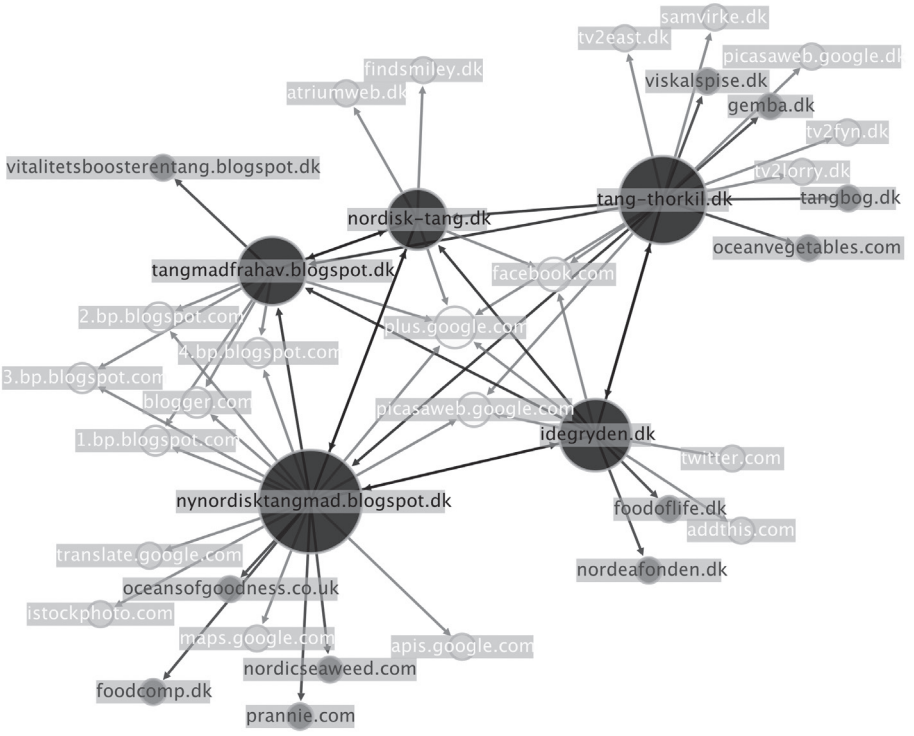
Marcus reasoned that multi-sited ethnography would nonetheless be necessary – inevitable, in fact – even if there was ground for methodological concern. He pragmatically asserted that much contemporary anthropological research was already *de facto* operating across multiple locales, and he identified a growing need for concrete "modes

of construction" to help ethnographers trace a space of research and map their object of study under such conditions. I want to suggest that the reappropriated use of a web crawler and a strategy for following hyperlinks through an emerging corpus of websites might fruitfully constitute one such mode (see also Munk, 2013).

The quali-quantitative as curation

If we decide on a strategy of following hyperlinks and use that strategy to trace the contours of an ethnographic field, then we quickly run into questions about the nature of such links. Indeed, hyperlinks themselves, as social phenomena, have not escaped ethnographic attention in the past (Beaulieu, 2005; Beaulieu & Simakova, 2006). Figure 3 shows a subset of the web corpus, namely a network of hyperlinked sites crawled from the profile page of Seaweed Thorkil. It makes it clear why careful consideration must still be given to the context in which links are provided and why the analysis of hyperlinks for this purpose cannot simply be automated but must retain a qualitative component. Besides the links shown on the screenshot of the page in Figure 4, we have to deal with links to ad services and web infrastructure and the fact that our tool, Navicrawler, is not capable of subdividing web entities like Twitter or Picasa into individual accounts (a feature that is now common in most crawlers).

Figure 3. *Curating prospect web entities for the corpus*



Comment: Starting from the user profile of Thorkild Degn Johansson on www.idegryden.dk, a network of five websites emerged (black nodes). They are all related either to his business as a seaweed retailer or to his advocacy of seaweed as an edible and sustainable resource. The network has been spatialized using a spring-based algorithm (ForceAtlas2) to make it possible to explore the next layer of potentially interesting neighbouring sites (grey nodes) and those sites that I had already decided, through Navicrawler, not to explore further (white nodes).

The concept of web crawling stems from the early days of the internet when search engines were not yet available and the only means of online navigation was to browse from one website to another. A crawler is essentially an algorithm that automates this way of exploring and mapping the web. It works from a seed list of known starting points from which it follows all the available hyperlinks and draws up a list of neighbouring sites. From this new list of starting points, it then reiterates the process until, in theory, it ends up having indexed every connected site on the web. In practice, however, crawlers cannot possibly keep up with the pace at which the web is evolving, so they are continuously at work (re-)indexing websites to make possible, for example, the now taken-for-granted operation of a search engine. However, web exploration with a crawler cannot be reduced to this perpetual, indiscriminate and automatic undertaking. It can be restricted, put on a leash, to focus deliberately on a region of the web, typically found in a specified link distance from a meaningful seed list. This has proven useful for websphere analysis (Schneider & Foot, 2005) or issue mapping (Marres 2015) and hinges on what Rogers (2009) called the "re-purposing" of online media, in this case websites and hyperlinks, for social research.

In this version of quali-quantitative work, the quantitative mapping of onlife traces needs qualitative curation to be meaningful. We are, as Rogers (2018) put it in his call for "critical analytics", obliged to think critically and reflexively about the way in which we appropriate natively digital traces. Tools such as crawlers or scrapers can be unleashed and run automatically, programmed to follow the shortest paths or the most-cited nodes, but these purely quantitative means of curation would not have been able to generate a New Nordic web corpus. This much is evident from Figure 3. By following hyperlinks indiscriminately and without further qualification from the profile page of Seaweed Thorkil, we would, in one iteration of the crawl, have been in the realm of Google, Picassa, Twitter, Blogspot and shortly thereafter the entire top layer of the web. Qualitative curation is necessary, and Navicrawler was designed specifically for this task.

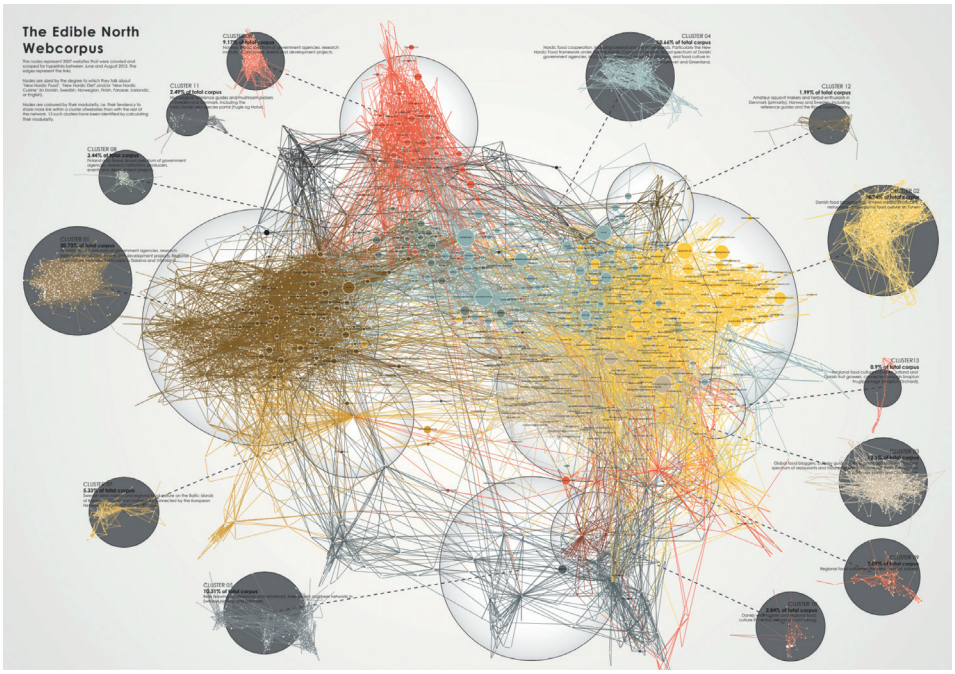
If we intend to use hyperlinks (or any other digital artefacts) as a metric for meaningful associations in relation to a theme like New Nordic Food, we cannot appropriate them uncritically. Here we can draw on the work undertaken at the interface between media studies and STS (e.g. Marres & Gerlitz, 2016). The quali-quantitative as critical metrics is about understanding that seemingly technical choices about algorithms and data tools are themselves worthy of social inquiry. What is a hyperlink, for instance? In what context? At what point in time? Those questions must be pursued in tandem with any effort to reappropriate hyperlinks to perform social research. In mixed-method research, this is often thought of in terms of sampling strategies. Probabilistic approaches to sampling, which in our case could be a decision to construct the web corpus by letting the crawler proceed *x* number of link steps from the initial seed, must be supplemented with more purposive approaches, in which we evaluate qualitatively which websites are relevant to our topic (e.g. Teddlie & Yu, 2007).

The quali-quantitative as algorithmic sensemaking

Finally, as a point of discussion, I would like to suggest a fourth possible interpretation of the hyphen in quali-quantitative work. It takes a radical approach to the computational turn and claims that some data-intensive methods, such as network analysis and espe-

cially algorithmic forms of community detection and pattern recognition, can themselves be understood as a form of qualitative reasoning that has noteworthy affinities with how ethnographers work in the field. Ethnography, at its core, is an explorative undertaking. It is, as James Spradley noted, as much about discovering relevant questions from informants as it is about discovering the answers to them (Spradley, 2016). Big data methods have often been critiqued for their lack of theoretical commitment. Similar critiques have been raised about ethnography, which can rarely claim to be testing hypotheses and is rather in the business of revealing how native theories order the world.

Figure 4. *Poster of the full new Nordic web corpus that the author carried with him on fieldwork*



Comment: The full new Nordic web corpus (Munk & Ellern, 2015) comprises 2007 food-related websites from Scandinavia that have all been scraped for hyperlinks using Navicrawler (Jacomy et al., 2007). It has been spatialized in Gephi with ForceAtlas2 (Jacomy et al., 2014). The nodes have been colour coded by their modularity (Blondel et al., 2009).

One of the most interesting outcomes of mapping the New Nordic Food phenomenon through hyperlinks was that it jeopardized my assumptions about relevant empirical differences in relation to gastronomy in Scandinavia. Since the topic was about geographical typicality, about landscapes and terroirs, it seemed reasonable that geography would also define the way in which actors associated online. Indeed, this turned out to be the case most of the time. I have for example explored above how a Norwegian cluster of websites talks about locally produced food. In other instances, however, it was not possible to label a cluster of websites by their nationality or regional affiliation. This was the case for beer brewers, mushroom pickers, wine producers, foragers and schnapps makers (some of the smaller clusters in Picture 2). The websites committed to these practices were more interested in associating with others that were interested in the same theme than they were in associating with others from the same area.

If I had discovered these differences in the association of food practices in the course of an interview or a stint of participant observation, I would have been very happy to count it as an ethnographic result. It is precisely these types of insights into native orderings of the world, which arise from following the actors on the ground, that constitute the objective of much ethnographic work. As Michael Agar put it: "Traditional social science is on the lookout for variables; ethnographers are on the lookout for patterns" (Agar, 2006: 109). However, in this particular instance, it was not the ethnographer who found those patterns. On the contrary, they were detected by running a spring-based layout algorithm on the network of websites comprised in the web corpus. ForceAtlas 2 (Jacomy et al., 2014) is designed to push nodes apart unless connected by edges, which then act as springs to hold the nodes together. The result is a visual landscape in which nodes are placed in relative proximity to other nodes to which they are densely connected and at a relative distance from nodes to which they are not connected. The algorithm is indeterminate, meaning that it never ends with exactly the same visual result. The same applies to the Louvain modularity algorithm used to colour the network in Picture 2 (Blondel et al., 2008; Lambiotte et al., 2008). This algorithm attempts to find the most optimal way of cutting the network into smaller components, severing as few edges as possible in the process. Again, there is no exact and reproducible result. It could be said that these types of algorithms work exploratively to find patterns in ways that are in fact rather similar to those that an ethnographer would use.

In mixed-methods research, Greene referred to explorative analysis as an area in which certain kinds of quantitative and qualitative analysis seem to have a lot of overlap: "in particular, across quantitative (postpositivist) and qualitative (interpretivist) studies. This strategy may yield 'stories that converge' or discrepancies that invoke fresh perspectives" (Greene et al., 1989: 257). Indeed, machine learning and data mining are, in general, ways of training algorithms to spot patterns. Although not compatible with conventional deductive statistics, it follows many of the principles laid down by, for example, John Tukey in his work on explorative data analysis (Tukey, 1977). Ethnography is in a similar predicament vis-à-vis its siblings in the social sciences. It is often in a situation in which it does not presume to know the theory of the world from which to generate hypotheses or ask questions and must thus first build a native worldview from the bottom up, exploratively and inductively.

Conclusion

I began this article with the frustrations of Bronislaw Malinowski, who had mapped all the observable aspects of culture in his Trobriand village without coming any closer to getting the "hang of tribal life". I compared that frustration to the meaning problem in digital methods, in which it is often not clear how we are supposed to interpret onlife traces, even if we can relatively easily accrue such data traces in large volumes. I have proposed four ways of mobilizing quali-quantitative responses to that problem. In the complementary style of quali-quantitative analysis, the meaning problem is both cast and solved as a qualitative matter. Again, following Tricia Wang (2013), numbers need stories! That position leaves little doubt as to the division of labour between the quantitative and the qualitative components of a piece of research. If we instead adopt a single level of analysis, however, this division of labour is immediately compromised. The

meaning problem is no longer about explaining the intentions behind onlife traces but about understanding how macro phenomena like web corpuses are enabled by interactions on the micro level. The way in which actors make associations can be followed in hand-held, qualitative ways or it can be tooled up and pursued quantitatively with a web crawler, but the basic *modus operandus* of the analysis remains essentially the same. If we adopt a curatorial approach to onlife traces, the qualitative component takes on a critically reflexive role in terms of understanding the media from which we harvest onlife traces for quantitative analysis and making choices about how to reappropriate such traces for social research. Finally, I have proposed that we might think fruitfully about algorithms for community detection and pattern recognition as performing essentially qualitative work. I have called this fourth style of quali-quantitative analysis *algorithmic sensemaking*, because it pushes some of the explorative, inductive work of the ethnographer to adopt various machine learning techniques.

Funding

This work was supported by the Carlsberg Foundation [grant number 2011_01_0811].

Note

1. Retrieved from https://www.norden.org/en/information/new-nordic-food-manifesto [accessed 2018, October].

References

Agar, M. (2006, September). An ethnography by any other name ... In *Forum Qualitative Sozialforschung/ Forum*, 7(4).

Appadurai, A. (1988). *The social life of things: Commodities in cultural perspective*. Cambridge University Press.

Barry, A. (2010). 12 Tarde's method. In M. Candea (ed.) *The social after Gabriel Tarde: Debates and assessments.* Vol. 4, (p. 177) Routledge.

Beaulieu, A. (2005). Sociable hyperlinks: An ethnographic approach to connectivity. In *Virtual methods: Issues in social research on the Internet* (pp. 183-198).

Beaulieu, A. & Simakova, E. (2006). Textured connectivity: An ethnographic approach to understanding the timescape of hyperlinks. *Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics*, 10: 6.

Blok, A. & Pedersen, M. A. (2014). Complementary social science? Quali-quantitative experiments in a big data world. *Big Data & Society*, 1(2): https://doi.org/10.1177/2053951714543908

Blondel, V. D., Guillaume, J. L., Lambiotte, R. & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 10: P10008.

Boellstorff, T., Nardi, B., Pearce, C. & Taylor, T. L. (2012). *Ethnography and virtual worlds: A handbook of method*. Princeton University Press. ISO 690.

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research*, 6(1): 97-113.

Callon, M. & Latour, B. (1981). Unscrewing the big Leviathan: How actors macro-structure reality and how sociologists help them to do so. In K. Knorr & A. Cicourel (ed.) *Advances in social theory and methodology: Toward an integration of micro-and macro-sociologies* (pp. 277-303) London: Routledge.

Chen, H. T. (2006). A theory-driven evaluation perspective on mixed methods research. *Research in the Schools*, 13(1): 75-83.

Creswell, J. W. (2014). *A concise introduction to mixed methods research*. London: Sage Publications.

Geertz, C. (2005). Deep play: Notes on the Balinese cockfight. *Daedalus*, 134(4): 56-86.

Girard, P., Dedinger, B., Ricci, D., Ooghe-Tabanou, B., Jacomy, M., Plique, G. & Tible, G. (2016). RICardo Project: Exploring 19th Century International Trade. In *Digital Humanities 2016: Conference Abstracts. Jagiellonian University & Pedagogical University, Kraków*, (pp. 208-210).

Greene, J. C., Caracelli, V. J. & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3): 255-274.

Jacomy, M., Ghitalla, F. & Diminescu, D. (2007). Méthodologies d'analyse de corpus en sciences humaines à l'aide du Navicrawler [Methods for corpus analysis in the humanities by way of the Navicrawler]. In *Programme TIC-Migrations.* Paris: Fondation de la Maison des Sciences de l'Homme.

Jacomy, M., Venturini, T., Heymann, S. & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PloS One*, *9*(6): e98679.

Johnson, R. B., Onwuegbuzie, A. J. & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2): 112-133.

Lambiotte, R., Delvenne, J. C. & Barahona, M. (2008). Laplacian dynamics and multiscale modular structure in networks. *ArXiv Preprint*: 0812.1770.

Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford: Oxford University Press.

Latour, B., Jensen, P., Venturini, T., Grauwin, S. & Boullier, D. (2012). "The whole is always smaller than its parts" – A digital test of Gabriel Tardes' monads. *British Journal of Sociology*, 63(4): 590-615.

Malinowski, B. (1922). *Argonauts of the Western Pacific: An account of native enterprise and adventure in the archipelagoes of Melanesian New Guinea*. London: George Routledge and Sons.

Marcus, G. E. (1995). Ethnography in/of the world system: The emergence of multi-sited ethnography. *Annual Review of Anthropology*, 24(1): 95-117.

Marres, N. (2015). Why map issues? On controversy analysis as a digital method. *Science, Technology, & Human Values*, 40(5): 655-686.

Marres, N. (2017). *Digital sociology: The reinvention of social research*. Cambridge: Polity Press.

Marres, N. & Gerlitz, C. (2016). Interface methods: Renegotiating relations between digital social research, STS and sociology. *Sociological Review*, 64(1): 21-46.

Munk, A. (2013). Techno-anthropology and the digital natives. In L. Botin & T. Børsen (eds), *What is techno-anthropology* (pp. 287-310).

Munk, A. K. & Ellern, A. B. (2015). Mapping the New Nordic issuescape: How to navigate a diffuse controversy with digital methods. In G. T. Jóhannesson, R. v d Duim & C. Ren, (eds.), *Tourism encounters and controversies: Ontological politics of tourism development* (pp.xx). Aalborg University Press

Munk, A. K. & Jensen, T. E. (2015). Revisiting the histories of mapping, *Ethnologia Europaea, Special issue: European Ethnology Revisited*, 44(2): 31.

Olwig, K. F. & Hastrup, K. (1997). *Siting culture: The shifting anthropological object*. Psychology Press.

Rogers, R. (2013). *Digital methods*. MIT Press.

Rogers, R. (2018). Digital traces in context| Otherwise engaged: Social media from vanity metrics to critical analytics. *International Journal of Communication*, 12: 23.

Schneider, S. M. & Foot, K. A. (2005). Web sphere analysis: An approach to studying online action. In Christine Hine (ed.), *Virtual methods: Issues in social research on the Internet* (pp. 157-170). New York: Berg Publishers.

Spradley, J. P. (2016). *The ethnographic interview*. Long Grove: Waveland Press.

Tarde, G. (2015). Monadology and sociology. In *The social after Gabriel Tarde* (pp. 53-87). Routledge.

Tashakkori, A. & Creswell, J. W. (2007). The new era of mixed methods. *Journal of Mixed Methods Research.* (3-7).

Teddlie, C. & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, 1(1): 77-100.

Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2).

Urry, J. (1972). "Notes and queries on anthropology" and the development of field methods in British anthropology, 1870-1920. *Proceedings of the Royal Anthropological Institute of Great Britain and Ireland*: 45-57.

Venturini, T., Jacomy, M., Meunier, A. & Latour, B. (2017). An unexpected journey: A few lessons from sciences Po médialab's experience. *Big Data & Society*, 4(2).

Wang, T. (2013). *Big data needs thick data* [blog post]. Retrieved from http://ethnographymatters.net/blog/2013/05/13/big-data-needs-thick-data/ [accessed 2018, March 3].