Diaspora and mapping methodologies: tracing transnational digital connections with 'mattering maps'

Alinejad, D.; Candidatu, L.; Mevsimler, M.; Minchilli, C.; Ponzanesi, S.; van der Vlist, F.N.

Diaspora and mapping methodologies: tracing transnational digital connections with 'mattering maps'

DONYA ALINEJAD,* LAURA CANDIDATU,* MELIS MEVSIMLER,*
CLAUDIA MINCHILLI,* SANDRA PONZANESI*
AND FERNANDO N. VAN DER VLIST[†]

*Department of Media and Culture Studies,
Utrecht University, the Netherlands
D.Alinejad@uu.nl / L.I.Candidatu@uu.nl
M.Mevsimler@uu.nl / C.Minchilli@uu.nl /
(corresponding author) S.Ponzanesi@uu.nl

[†]DFG Collaborative Research Centre 1187 'Media of Cooperation',
University of Siegen, Germany / Department of Media Studies,
University of Amsterdam, the Netherlands,
fernando.vandervlist@uni-siegen.de

Abstract There is a methodological tendency in work on diaspora and digital media for quantitative investigations to approach diaspora in static ways that contrast with theories of diaspora as a dynamic cultural formation. On the other hand, qualitative, ethnographic work tends not to engage with digital methods and quantitative data-driven investigation. In this article, we sketch this methodological and disciplinary disconnect and address it by proposing a model for understanding digitally mediated formations of diaspora that combines digital methods techniques with a sensitivity to ethical and theoretical discussions of migration and diaspora. Drawing on interpretive epistemologies and feminist research ethics, we present a case study analysis of a locally informed, Turkish–Dutch issue. We argue for a method that produces 'mattering maps'. This involves tracking and visualizing digital traces of an issue across web platforms (Google Search results, Facebook pages, and Instagram posts) and integrating this with an analysis of the face-to-face interview responses of a key issue actor.

In late 2015, a Turkish–Dutch girl named Bade Çakır was diagnosed with acute childhood lymphocytic leukaemia. In January 2016, the four-year-old girl's parents, who lived with Bade in Rotterdam, the Netherlands, made a public call for stem cell donations. They specifically addressed the call to members of the Turkish community in the hope that widening the pool of Turkish donors would increase the chances of finding a match. The call for donors originated from the family and spread online through multiple social media channels; the case also received local and national news coverage in the Netherlands and Turkey. Despite these efforts, Bade's condition rapidly worsened in late February and she died that same month. We learned of this case in March 2016 through a third-generation member of the Turkish–Dutch community, who did not know Bade or the family, but spoke to one of our team members in an interview about the tragic story and the significance of the collective mobilization. This prompted us to initiate a more targeted online search on this issue.

What initially struck us about this case was that the urgent media mobilization and moving story failed to appear in our preliminary online investigations of the Turkish–Dutch diaspora. In line with previous approaches to the large-scale study of a diasporic web presence, we had originally started by compiling a corpus of Turkish diaspora websites for large-scale quantitative digital analysis, while conducting interviews for separate qualitative analysis. The invisibility of this apparently significant event in the initial review of websites brought into sharp focus the problem of how and where to study a diaspora group online and, by extension, how more effectively to integrate qualitative and quantitative methods in doing so. The literature on migrant group formations and digital media reflects various approaches to the problem of locating diaspora online, but there is a relative lack of work that incorporates these different approaches, conceptually and methodologically. Typically, 'e-diasporas' (Diminescu et al. 2011) are studied in communication studies with a focus on web hyperlinks and connectivities; 'digital diasporas' (Brinkerhoff 2009) are studied in migration studies through discourse analysis, primarily looking at blogs, forums and websites; and in ethnographic media studies the tendency is to study migrants' web uses as embedded in their everyday practices (Gomes 2016; McKay 2012; Madianou and Miller 2012).

Our experience with Bade's case raised the practical problem of which websites to include in a corpus of meaningful web-based social phenomena – however temporally bounded – in our analysis. The issue's temporality and emotional significance also raised the conceptual question of how to incorporate the dynamism of diasporic connections into an analysis of web-based connectivity. The seminal works of such scholars as Hall (2003) and Gilroy (1993) have demonstrated the dynamic, context-based nature of identity formation processes and argued for the hybrid character of diasporic identity. Others have compellingly pointed out how static approaches to diaspora neglect the diversity of diasporic experiences in their racialized, classed and gendered outcomes (Anthias 1998; Brah 1996; Clifford 1994). Hence, developing digital research approaches that pre-define a corpus of migrant/diaspora themed websites as the relevant ones runs the risk of reifying and/or overstating the importance of ethnic and national categories, thus reproducing 'methodological nationalism'

(Wimmer and Glick Schiller 2002) by seeing migrants as predominantly preoccupied with national/ethnic identifications.

In this article, we argue in favour of foregrounding the relationship between theory and empirical data when discussing new methodological approaches to the digital study of diaspora, a plea that gains urgency in the current context where 'big data' driven approaches to migrant web use are beginning to acquire institutional dominance. Rather than succumb to 'digital positivism' (Fuchs 2017), our aim is to develop useful integrations of newer digital methods within qualitative research epistemologies applied to diaspora research, paying attention to the implications of how to theorize diaspora formation in contemporary contexts of digital media proliferation. To this end, in the section below, we discuss how we operationalize digital diaspora. Then, in the section on 'Mapping Bade Çakır and the Turkish diaspora online', we lay out our empirical comparative analysis across three main platforms. In 'The social ethics of mapping', we consider the complex ethical implications of 'big data' research that surface using specific research tools. In the section on 'Integrating digital methods with data ethnography', we draw on ethnographic research principles to acknowledge the partiality and situated nature of the mapping exercise itself and contextualize the meanings of the mapped connections. Ultimately, we suggest a method for creating diasporic 'mattering maps', on which we elaborate with reference to a mixed-methods investigation of Bade Çakır's case through the analysis of digital data and an interview with a key stakeholder and actor (namely Bade's father, Muhammet Çakır).

Operationalizing digital diaspora

The application of traditional ethnographic methods to online or 'virtual' environments is currently the subject of debate, and the notions of 'virtual ethnography', 'digital ethnography', and 'netnography' have emerged to refer to a range of specialized qualitative research techniques and epistemologies (Hine 2015; Kozinets 2009; Pink et al. 2016). The recent ethnographic work on digital media use has also included quantitative survey data about digital media use alongside ethnographic insights (see Johnson and McKay 2011; Miller et al. 2016), but Diminescu and Loveluck (2015) remark that importing research methodologies suitable for the 'offline' world to the web does not account for the web's main feature of 'connectivity'. Despite their diversity – including ethnographic work (Georgiou 2006; Nakamura and Chow-White 2012) and discourse analysis (Trandafoiu 2013) – qualitative approaches have tended not to engage with the proliferating methodological possibilities presented by tools and approaches oriented towards large-scale, quantitative, digital data-based research, and 'digital methods' (Rogers 2013). Digital methods are oriented towards the study of social phenomena that are specific to, and inseparable from, the devices or platforms on which they emerge (for example, links, engines, likes, shares, tweets), which involve the use of research methods geared towards the study of 'natively digital' objects, devices, and constitute a relatively new field of research.

Such digital methods approaches have been successful in producing digital representations of national and religious diasporas' online presences in what appear to be

comprehensive visual records; one example is Kok and Rogers's (2016) work on the Somali diaspora. However, these visualizations offer a series of snapshots that produce partial, static visual accounts of a socio-cultural phenomenon that constructivist theorizations of diaspora have defined as a continuous and layered formation process. Therefore, the temporal and emotional dimensions of connectivity that make connections dynamic and meaningful to people remain secondary. Consequently, in problematizing how digital methods have been applied to the study of diaspora thus far, we raise the core question of whether what is being mapped is, in fact, what matters. What is the significance of hyperlinks between migrant websites and/or connections between social media profiles and pages? And what place do they have within migrants' lived realities? We call for a mode of mapping that seeks to understand the social meanings of digital connections for users.

For this reason, we use an 'issue mapping' approach, taking Rogers et al. (2015) as a starting point. Issue mapping is a way of exploring the utility of online data and devices to trace how a particular topic of discussion mobilizes conversation and contestation through the aggregate engagements of distinct users. In this sense, it is suitable for investigating connections as they are produced by the connecting practices of users who are moved to engage with a particular issue in a specific way. Issue mapping methodologies were originally developed to detect and track the emergence of 'current affairs'. However, scholars employing issue mapping have also discussed the problem of defining an issue. For instance, Marres (2015) has pointed out tensions between what involved actors and issue experts define as an issue (for example, based on in-depth knowledge) and what the platform defines as an issue (for example, based on top trends). This problem leads her to argue that 'it won't do for issue mapping research to call an "issue" whatever the platform says is one' (Marres 2015). By taking users' lives as a starting point, as we tried to do by choosing this case, we see how smaller-scale issues emerge and come to matter to people in intensely significant ways. In this sense, we treat migrant research respondents or web users themselves as the experts when it comes to identifying the topics that engage them.

In general, we are 'repurposing' the affordances of digital devices (Weltevrede 2016) – such as data collection tools typically used for large-scale digital research – to map diasporic connections. We then bring these tools together with Grossberg's (2010) influential discussion of 'mattering maps' to map how people within the Turkish–Dutch community were moved to connect to Bade Çakır's case. Grossberg uses this notion to theorize the affective dimension of how certain issues come to take on intensity for people, thus moving them to action. For him, such affective traces order our experience, and through them certain differences of identity come to matter (more than others). As he argues, 'the image of mattering maps points to the constant attempt, whether or not it is successful, to organize moments of stable identity, sites at which we can, at least temporarily, find ourselves "at home" with what we care about' (Grossberg 2010).

The idea of mattering maps inspires us in two senses. First, the affective layer within migrants' digital media practices gives us a way of focusing on how digital connections are imbued with investment and therefore social and emotional significance, namely how something comes to *matter* (mattering as a verb). Second, it allows us to bring this

focus on affectivity to existing digital methods that are useful for mapping online data, namely taking the mapping of a *matter* of concern (for example, a question or controversy espoused, advocated, and upheld by specific actors or sub-groups) as a starting point, and thus avoiding methodologically reifying ethnic or national categories. To concretize the rather metaphorical 'image' of 'mattering maps' that Grossberg proposes, we combine elements of digital methods with the emergent 'data ethnography' approach (Dourich et al. 2016) proposed by anthropologists and scholars from other, typically qualitative, fields to develop ways to map empirical traces of an issue.

By tracing the issue's formation and movement across different online platforms and pages, we observed distinct 'spheres' – 'device-demarcated source sets' (Rogers 2013) – across which the issue emerged. First, we observed an institutional, news-related sphere (via Google Search); second, a cause-related sphere (via Facebook pages); and third, an image sphere (via Instagram and Google Images). We take each of these spheres in turn in the following sections, after which we compare and reflect on some of the ethical questions this research raised for us. From these reflections on the problems of representing 'the other' and conducting research on human participants, the last step of the analysis emerged – a digital ethnographic account based on a semi-structured interview. We analysed this interview through an initial open-coding step (in which four of the team members independently open-coded the two-hour long interview), after which we developed the shared code tree used to perform the final coding and produce the textual analysis.

Mapping Bade Çakır and the Turkish diaspora online

Google's news sphere

The first step of our investigation involved analysing Google Search results across local Google domains (Rogers 2013) for the literal query ['Bade Çakır']. We explored three local domains – *google.nl* (the Netherlands), *google.com.tr* (Turkey), and *google.de* (Germany), and compared the top-ranking web results for each domain (Figure 1).[1] We subsequently compared the three distinct result lists to identify differences (namely unique URLs appearing in only one result list) and commonalities (namely URLs shared by at least two result lists). In addition, we collected publication dates and identified languages for each article. Regarding the strong overlaps between the Turkish and German domains, we quickly noticed that there was a recurrence of hyperlinks from the *.de* domain in the Turkish results, which made us aware of the strong connection with the Turkish–German diaspora.

When we viewed the websites to which the listed hyperlinks led, it was immediately apparent that the Google results were almost exclusively news items. Hence, we saw the way Google gathered these items as the effect of the device generating a news-related sphere around this issue, a finding we discuss below with a comparison across platforms. Furthermore, the comparison reveals that the highest number of unique hyperlinks belongs to the Dutch domain list and is published in Dutch, while the Turkish and German domain lists have a high degree of overlap and are both mainly published in Turkish. This suggests that initially the issue's circulation was largely

contained within the Netherlands – the place where it originated – and concerned the Turkish–Dutch community. This is also evidenced in the finding that when the dates are considered, most of the articles on the *.nl* pages were posted from 23–26 January 2016, when the call for stem cell donation was distributed. This contrasts with the posts from the Turkish and German domains, which had the most hyperlinks in common and were posted about a month later, from 24–26 February 2016, when Bade died. Notably, the majority of the hyperlinks that appear in the German domain are *.com.tr* hyperlinks, which further suggests that the issue ended up in the German diaspora via Turkey and the Turkish-language web.

Figure 1: Triangulation tables for Google Search results for ['Bade Çakır']

Google Search – Web	Triangulate: Lists colored by frequency, ordered by frequency, original order							
Frequency	1			2			3	
Language	nl			tr			de	
Country	The Netherlands			Turkey			Germany	
Google domains (ccTLD)	google.nl			google.com.tr			google.de	
Google Search market share by country (Austin Return On Now Internet Marketing	96%			94%			94%	
Date Captured	09/04/2016			09/04/2016			09/04/2016	
Result rank	Result URL	Article date	Langu	Result URL	Article date	Langu	Result URL	Article dat Langu
1	rijnmond.nl/nieuws/139054	24/02/2016	nl	haberler.com/bade-bebek-s	25/02/2016	tr	haberler.com/bade-bebek-s 25/02/2016 tr
2	rijnmond.nl/nieuws/137508	23/01/2016	nl	milliyet.com.tr/minik-bade-s	26/02/2016	tr	youtube.com/watch?v=lu_yi 25/02/2016 tr
3	youtube.com/watch?v=xjbtf	25/01/2016	nl	milliyet.com.tr/kemik-iligi-ai	24/02/2016	tr	youtube.com/watch?v=1zxq 26/03/2016 tr
4	metronieuws.nl/nieuws/rott	27/01/2016	nl	sabah.com.tr/yasam/2016/0	25/02/2016	tr	milliyet.com.tr/minik-bade-s 26/02/2016 tr
5	metronieuws.nl/nieuws/rott	23/01/2016	nl	haberturk.com/saglik/haber	24/02/2016	tr	haberturk.com/saglik/haber 26/02/2016 tr
6	facebook.com/rtvrijnmond/	23/01/2016	nl	hurriyet.com.tr/badeden-aci	24/02/2016	tr	haberturk.com/spor/futbol/ 15/02/2016 tr
7	ed.nl/regio/eindhoven/bade	05/02/2016	nl	fanatik.com.tr/2016/02/24/t	24/02/2016	tr	sabah.com.tr/yasam/2016/0 25/02/2016 tr
8	puntuit.nl/nieuws/vechten-	11/02/2016	nl	posta.com.tr/saglik/haberde	24/02/2016	tr	yenisafak.com/hayat/minik- 24/02/2016 tr
9	turksemedia.nl/hoog-bezoel	27/01/2016	nl	youtube.com/watch?v=lu_yi	25/02/2016	tr	hurriyet.com.tr/badeden-aci 24/02/2016 tr
10	haberler.com/bade-bebek-s	25/02/2016	tr	yenisafak.com/hayat/minik-	11/02/2016	tr	fanatik.com.tr/2016/02/24/t 24/02/2016 tr

Google Search – Images	Triangulate: Lists colored by frequency, ordered by frequency, original order							
Frequency	1			2			3	
Language	nl			tr			de	
Country	The Netherlands			Turkey			Germany	
Google domains (ccTLD)	images.google.nl			images.google.com.tr			images.google.de	
Google Search market share by country (Austin Return On Now Internet Marketing	96%			94%			94%	
Date Captured	30/01/2017			30/01/2017			30/01/2017	
Result rank	Result URL (page)	Article date	Langu	Result URL (page)	Article date	Langu	Result URL (page)	Article dat Langu
1	rijnmond.nl/nieuws/137508	23/01/2016	nl	61saat.com/bolgesel/trabzo	31/01/2016	tr	61saat.com/bolgesel/trabzo 31/01/2016 tr
2	youtube.com/watch?v=xjbtf	25/01/2016	nl	yenivatan.be/losemi-hastasi	23/01/2016	tr	yenivatan.be/losemi-hastasi 23/01/2016 tr
3	lokum.nl/ziek-meisje-4-zoek	23/01/2016	nl	fanatik.com.tr/2016/02/24/t	24/02/2016	tr	fanatik.com.tr/2016/02/24/t 24/02/2016 tr
4	turksemedia.nl/hoog-bezoel	27/01/2016	nl	zamanhollanda.nl/haber/bo	25/02/2016	tr	sonhaber.eu/gundem/bader
5	rijnmond.nl/nieuws/137508	23/01/2016	nl	yenisafak.com/gundem/bad	28/01/2016	tr	linetv.com.tr/tag/bade-cakir tr
6	puntuit.nl/nieuws/vechten-	11/02/2016	nl	sonhaber.eu/gundem/bader		tr	zamanhollanda.nl/haber/bo 25/02/2016 tr
7	libelle.nl/mensen/actueel/ni	28/01/2016	nl	linetv.com.tr/tag/bade-cakir		tr	yenisafak.com/gundem/bad 28/01/2016 tr
8	agf.nl/artikel/136861/Fruitin	12/02/2016	nl	sonhaber.eu/gundem/bade-	31/01/2016	tr	sonhaber.eu/gundem/bade- 31/01/2016 tr
9	rijnmond.nl/nieuws/137508	23/01/2016	nl	milliyet.com.tr/kemik-iligi-ai	24/02/2016	tr	youtube.com/watch?v=xjbtf 25/01/2016 nl
10	turksemedia.nl/vrouw-turks	11/02/2016	nl	hurriyet.com.tr/badeden-aci	24/02/2016	tr	rijnmond.nl/nieuws/137508 23/01/2016 nl

Note: search term queried in several local Google domains: google.nl, google.com.tr and google.de (upper section: top ten web results as from 9 April 2016; lower section: top ten image results as on 30 January 2017). Colour-coding is based on the degree of commonality across lists (shades of blue) and based on article-level language detection (Dutch, Turkish, or German).

Overall, this shows that the circulation of the issue of Bade Çakır on Google was characterized by its emergence in the Netherlands and its spread across the Dutch news sphere, its subsequent movement to Turkey a month later through Turkish-language news coverage, and finally to Germany where the large German-based Turkish diaspora

engaged with the issue in Turkish. Given the nature of this circulation, we see how the issue engaged the Turkish diaspora differently in different national settings (German and Dutch), with Turkey and the Turkish language playing a key role in mediating these cross-diasporic connections.

Facebook's cause sphere

We noticed that one of the hyperlinks obtained via Google made reference to an organization that had supported the parents' call for donations.[2] The organization is called Turkish Volunteers Foundation,[3] and although they did not have a website, they had a Facebook page titled 'Hollanda Türk Gönüllüler Vakfı',[4] which we later learned constituted the sole online platform on which this organization actively disseminated content relating to Bade's case. This suggested to us that Facebook was a potentially significant social media platform in the spread of the issue, and the page of this Turkish–Dutch organization gave us an entry point to start investigating how this issue spread on this platform. Through this investigation we came across another page. Initially named 'Bade İçin Bir İyilik, Bir İlik…',[5] this second page was later renamed 'Bade Foundation', after the foundation set up in Bade's memory.[6]

Using Netvizz (Rieder 2013), we were able to extract data from Facebook pages. This enabled us to explore page-like networks around the combined page-like networks of 'Bade Foundation' and 'Hollanda Türk Gönüllüler Vakfı' (Figure 2). This revealed strong connections to 17 other Facebook pages, where a connection is constituted by an action of 'liking' (in at least one direction) between pages. Notably, neither of the pages states any address or geographical location in the page's 'About' section. However, by tracing direct connections with other pages it is possible to derive their locations from the set of closely related pages in their network. This location proxy method shows a clear embedding of this Turkish–Dutch issue within Rotterdam. The use of language in descriptions, posts, and comments on these pages supports this hypothesis (namely they are mainly in Dutch or a combination of Dutch and Turkish). This network thus reflects a strong localization of the issue in and around the city of Rotterdam. That the pages in this network represent local organizations suggests that the circulation of the issue relied on local community-based involvement.

The network for 'Bade Foundation' is notably larger than the other network when taking into consideration directly related pages as well as indirectly related pages (namely crawl depth: 2; 1,441 nodes), and is heavily centred on the theme of sick children in the sense that most of the pages are concerned with particular cases of sick children and the work of various humanitarian NGOs dealing with the wider issue of childhood illness. This extended network contains many large humanitarian organizations, which are connected to the seed pages only through 'Tied together – bracelets for you and hospitalized kids',[7] a page managed by a young Dutch girl who sends free personalized bracelets to children with life-threatening illnesses. This explains the connection with Bade and the Netherlands but shows that the case of Bade Çakır loses its specific diasporic and ethnic dimension in this network as it gets lost in the broader category of humanitarian causes. The international circulation of the issue via the 'Bade

Foundation' page is therefore mediated by connections built on what we can call humanitarian ties that connect this issue to a wider cause sphere in the Netherlands and beyond.

Figure 2: Page–like networks for 'Hollanda Türk Gönüllüler Vakfı' (@htgvv) and 'Bade Foundation' (@badefoundation)

Depth	Title	Username	URL	Category	users_can_po	talking_abou	post_activity	fan_count	Languages		Region
0	Bade Foundation	@badefoundation	https://www.faceboo	Cause	yes	4827	0.03	19594	nl	tr	
0	Hollanda Türk Gönü	@htgvv	https://www.faceboo	Non-Profit Organizati	yes	1048	0.03	6909	tr		
1	Atv Avrupa Vizyon N	@vizyonnl	https://www.faceboo	TV Show	no	42	0.01	1489	tr		The Netherlands
1	DENK	@DenkNL	https://www.faceboo	Political Party	no	8669	0.08	64368	nl		The Netherlands
1	Diva Party Centre	@divapartycentre	https://www.faceboo	Performance & Sport	no	1102	0.01	22290	nl		Rotterdam Area
1	Dyna Wedding FX	@dynaweddingfx	https://www.faceboo	Company	yes	90	0	739	nl		The Netherlands
1	Fancy Palace	@fancypalace	https://www.faceboo	Beauty	yes	207	0.01	4333	nl		Rotterdam Area
1	Grup SILA NL	@grup.sila.nl	https://www.faceboo	Musician/Band	yes	14	0.02	5878	nl	tr	The Netherlands
1	Meld Islamofobie	@MeldpuntIslamofob	https://www.faceboo	Community	no	463	0.05	11394	nl		The Netherlands
1	MeloenFotografie	@meloenfotografie.n	https://www.faceboo	Camera/Photo	yes	34	0.01	1181	nl		Rotterdam Area
1	OBA Grandcafe en R	@obagrandcafe	https://www.faceboo	Local Business	yes	401	0.08	6269	nl		Rotterdam Area
1	Öztürk Foto Video	@ozturkfoto	https://www.faceboo	Media/News Compar	yes	151	0.01	11281	tr	nl	Rotterdam Area
1	SonHaber.eu	@SonHaberEu	https://www.faceboo	Media/News Compar	no	5605	0.72	21636	tr		
1	Stichting Matchis	@stichtingmatchis	https://www.faceboo	Other	yes	1036	0.01	12688	nl		The Netherlands
1	Taart Paradijs	@taartparadijs	https://www.faceboo	Food & Beverage Con	yes	58	0.01	4738	nl		Rotterdam Area
1	Thermen 1001		https://www.faceboo	Health/Beauty	yes	22	0	4247	nl		Rotterdam Area
1	Tied Together - Brac	@tiedtogetherbracel	https://www.faceboo	Non-Profit Organizati	yes	1	0.01	1405	nl	en	The Netherlands

Data collection: Netvizz v1.41; module: 'page like network'; settings: input Page IDs: 752828048195495, 1546625252320085; crawl depth: 1.

*Instagram's and Google's image spheres*

To expand further our analysis of the issue's spread across spheres and platforms, we surveyed other social media platforms for traces of Bade's case. Instagram emerged from this initial survey as a platform where the issue was spread in a different form than on Facebook, namely through hashtags and with a strong emphasis on photographs. This made Instagram an interesting point of comparison with the other platforms. Instagram data could be gathered using DMI Instagram Scraper, built on top of Instagram's API, which was able to retrieve either the latest media tagged with a specified term or the media around a particular location.[8] The extracted Instagram data cover the three-month period from late December to early April. Our analysis of Instagram is therefore based on the dataset we generated through this tool before the closing of its APIs in June 2016, consisting of 374 photographs and videos (we will use the term 'image' to refer to both) along with additional metadata, including the number of likes and comments, user profile information, tags and filters used.

Among other things we conducted a co-hashtag analysis to explore pairs of hashtags (Figure 3a–b). Most notably, hashtags that co-occur with #badecakir on Instagram are linguistically diverse, crossing over between Turkish, German, English and Dutch (Figure 3a). Although this diagram displays the overall distribution of hashtags proportionally across languages, it does not give insight into specific language crossovers. To identify language patterns, we therefore created associational profiles based on hashtags and their languages (Figure 3b). Each hashtag–language associational profile is composed of all associations in the source language (for example, all associations between co-occurring Turkish hashtags) plus all direct associations with co-occurring

hashtags in other languages (for example, all non-Turkish hashtags co-occurring with Turkish hashtags). Each node in Figure 3b represents a hashtag used in a post together with the seed hashtag (namely #badecakir) and is coloured according to the language of the co-occurring hashtag. Each connection represents a hashtag co-occurrence within the same Instagram post, and connections grow thicker as the number of co-occurrences increases. Connections are colour-coded according to the language of the co-occurring hashtag (for example, for the Turkish graph: TR-to-TR are red lines, TR-to-DE are black, TR-to-EN are blue, and TR-to-NL are orange).

This categorization of hashtags according to language illustrates the overall dominance of Turkish and the relative absence of Dutch in this sphere. This is something we see clearly in Figure 3a, which shows the proportionate prevalence of each hashtag and the language to which it belongs in the form of a tree diagram. Yet, most interestingly, Figure 3b shows repeated language crossovers within the same posts (namely multilingual Instagram posts). Hence, although we see that Turkish is by far the most prominent hashtag language (813 edges), the crossovers with German (180), English (136), and Dutch (19) within the same posts indicate how prevalent bilingual and/or multilingual posts are.

Figure 3a: Hierarchical treemap diagram for hashtags associated at least twice with the source query [badecakir] on Instagram on 19 April 2016
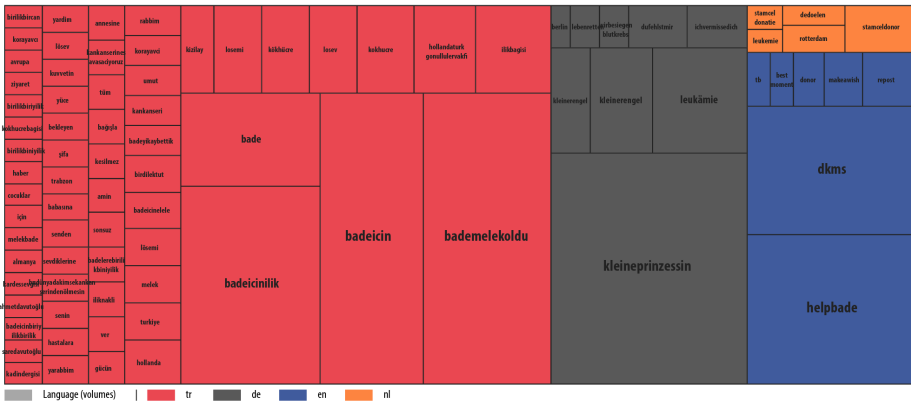
Hierarchy: language and hashtags; colour-coding: by language category; scaling: by frequency count. Data collection: DMI Instagram Scraper; method: tag [badecakir]; iterations: 1.

These findings suggest that the issue specifically addresses a Turkish-speaking community: hashtags such as '#badeicinilik',[9] '#bademelekoldu',[10] and '#kokhucre'[11] were circulated most. It is also noteworthy that the second most frequently used co-hashtag in the network was the German-language hashtag, '#kleineprinzessin'.[12] In addition, German was the second-most prominent language after Turkish, as there were 156 posts (out of 374) with German text and/or German hashtags. Further investigation of the network revealed that this issue's German sphere on Instagram was largely dominated by one account, which was set up specifically for Bade, making the issue more

visible for a German-speaking audience.[13] Interestingly, this account for Bade no longer exists so it may have been set up purely to spread the hashtag in Germany. The prominence of the Turkish language suggests that Instagram users were circulating Turkish hashtags to reach Turkish community members both within and outside the Netherlands to mobilize support for donating stem cells, and therefore chose to use Turkish as a means of reaching the target group. This is corroborated by the fact that Dutch is the least-used language within this network.

Figure 3b: Associational hashtag–language profiles (visual tagnets) for hashtags and languages associated at least twice with the source query [badecakir] on Instagram on 19 April 2016
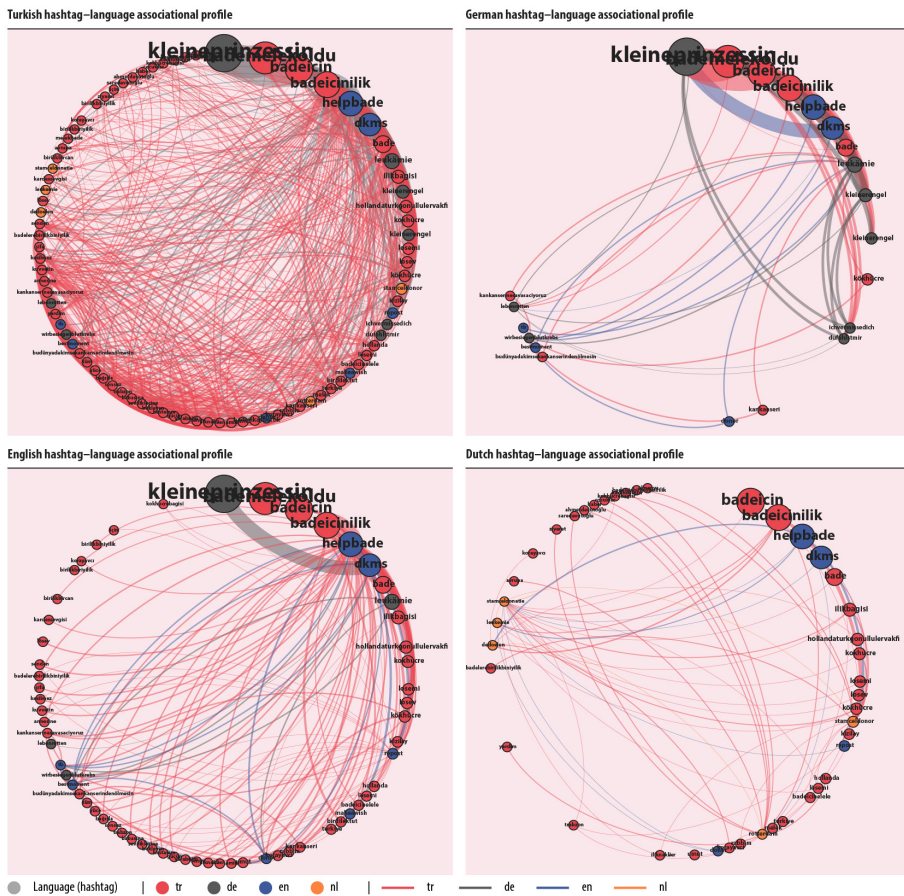
Nodes: 237; edges: 2,305. Filters: all nodes with fewer than two connections (63.71 per cent). Node scaling: by frequency count; colour-coding: by language category; layout: 'Circle Layout'; node layout direction: clockwise.

Furthermore, we noticed that some images and captions included the names and websites of stem cell banks in Germany, Belgium, and Austria, calling for members of

the Turkish diaspora in these countries to donate. While Instagram allows post geo-tagging, only 2 per cent of the posts in our dataset were geo-tagged, giving us little location information with which to work. However, relevant locations were referred to using hashtags in various languages, including the Netherlands ('Hollanda'), Germany ('Almanya'), Turkey ('Turkiye') and Rotterdam. Using language patterns to trace an issue thus helps understand how an issue evolves and travels across different linguistically defined social spaces, potentially cross-cutting national or regional borders and extending seamlessly to include diaspora web users across geographic locations.

Our temporal analysis of the Instagram data showed that the peak month for all posts using '#badecakir' is February 2016.[14] Closer investigation of some of these images revealed that their captions concern Bade's worsening condition. Although Bade died in late February, user engagement around the hashtag and photos of the girl continued to circulate on Instagram in the months that followed, albeit to a lesser degree. However, we saw the number of images around the hashtag increase after May 2016, from 374 to 871.[15] It is likely that the main reason for this is the continued effort to raise awareness and increase stem cell donation from Turkish people in the Netherlands and other countries. This is the objective of the aforementioned Bade Foundation,[16] which Bade's father, Muhammet Çakır, set up in August 2016. Posts linking to news articles about the foundation's success continue to circulate with titles such as 'Bade saved two lives',[17] referring to the social impact and legacy of the case.

Comparing across platforms

Taken together, these findings suggest that the purposeful international mobilization of the Turkish diaspora characterized the spread of the issue on Instagram and that this feature distinguishes its spread on Instagram from its circulation on Facebook pages. The language-based Instagram analysis reveals an interesting dimension of how this issue mobilizes community and thus how one can potentially trace diasporas on and across various web platforms. In turn, we saw a contrast between the issue's circulation on Facebook and the Google-related findings, since it did not have a cross-diasporic dimension (namely not reaching Germany) on Facebook, focusing rather on local Turkish–Dutch institutions. One way of visually summarizing the issue's spread across all the above-mentioned platforms is through a chronological horizon graph diagram of engagement per type and per platform (Figure 4a–b). This primarily shows that engagement with the issue across the three main platforms had different starting points and similar peaks (especially 24–26 February), and drew a diversity of languages, including many emoji-only comment responses. This not only suggests the different uses and audiences of the different platforms regarding this transnational issue (Figure 5), but also demonstrates the importance of tracing which media platforms diaspora actors (for example, individuals and institutions) take up and for which purposes. Rather than investigating an online diasporic presence by looking at web connectivity as an undifferentiated whole, tracing diasporic issue engagement across the most relevant platforms helps acknowledge how users who shaped how the issue spread took up each platform. Furthermore, analysing the contrasting results from Google alongside social

media platforms in a mixed methodological model reminds us that today's web does not only consist of social media platforms. This raises questions about the possible sociality of search engines, a platform type for which digital methods research tools have been developed but that remain vastly understudied by ethnographers interested in social network dynamics.

Figure 4a–4b: A multi-series horizon graph diagram (small multiples) combining time series data from multiple source platforms about Bade Çakır, allowing easier comparison with one another and with the annotated timeline diagram (Figure 4b)
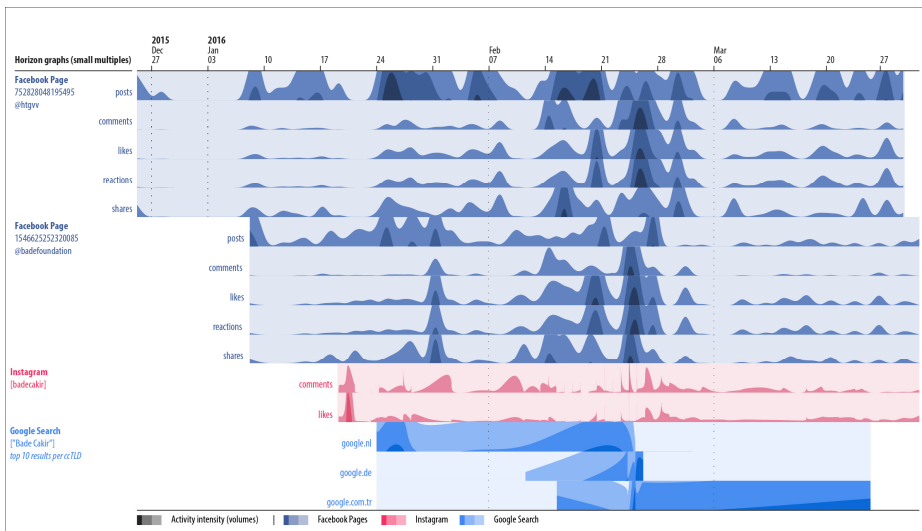

Grouping: source platform and data type (Facebook pages: user activity around posts on pages; Instagram: media tagged with a specified term; and Google Search: top ten results per ccTLD). Colour-coding: by source platform and data type. Scaling: by frequency count (interpolation: basis spline; mode: offset).

In addition to the digital traces mapped thus far, we undertook a creative, mixed-methods, comparative analysis of the Instagram images. We plotted the Instagram images according to their formal visual characteristics, using a plot of the top Google images of Bade as a point of comparison (Figure 6). These visualizations reflect a uniformity of clusters across the two platforms, which suggests the existence of an image sphere across both visually-oriented platforms, and the widespread reposting of the same or similar images. However, these visualizations also demonstrate an important challenge that visual content and image-based platforms present for quantitative digital analysis, one that remains unresolved by the large-scale overviews these visualizations afford: the overview produced (itself partial) in fact obscures what is visible from looking at the images themselves, and does not encompass the experience of being repeatedly exposed to (versions of) the same digital (audio) visual representations of a visibly ill four-year-old girl.[18] Although ethnographic research on the use

of Facebook photos has taken a mixed methods approach to understanding how social media photography genres fit into everyday social practices in specific cultural contexts (Miller and Sinanan 2017), studying how a single image circulates online poses a different methodological challenge, one that has only begun to be addressed through the technical capabilities of digital methods (Vis 2015).

Figure 5: A multi-series streamgraph diagram (small multiples) combining time series data from multiple source platforms about Bade Çakır showing the relative distribution of distinct languages used on each platform.
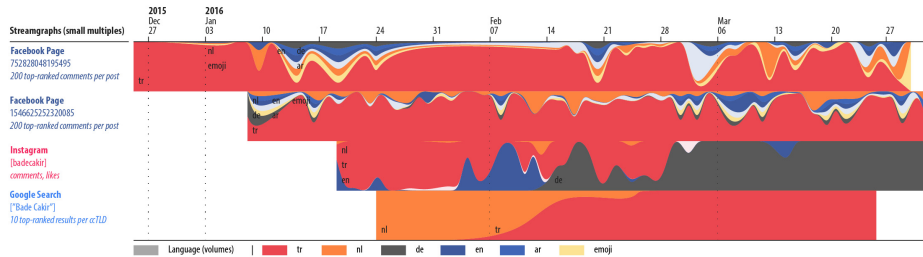

Note: Graphs are aggregated by source platform (namely Facebook pages, Instagram and Google Search) and sorted by chronological order. Colour-coding: by language category; by scaling; by language volume.

From the impressions we gathered from visually surveying the images circulated most intensively on Instagram, it became clear that the photos and videos were predominantly of Bade, featuring her as the main figure at various stages of her illness, mostly in the hospital setting. Dutch users in the network (some of whom we found to be active members of the foundation) circulated most of the images; these included audio-visual material produced or edited by users to explain the process of stem cell donation and to call on people to donate, using an informative, humanitarian tone. Most of the videos show Bade lying in her bed in a hospital room, sometimes with her parents or other visitors by her side; she looks directly at the camera as she thanks everyone who has tried to help her. This audiovisual record of an ill four-year-old girl raises sentiments and transmits affect, shaping the engagement between images and viewers (Gonzalez-Polledo and Tarr 2016).

Looking at these images and watching the videos showed us, more than any part of the analysis discussed so far, the intensely emotional dimension of this issue, which was also the most difficult to capture with the visualizations of digital data traces. This point signals a need to pay more attention in the work on digital diaspora to how to study emotional transmission empirically via web platforms as a mode of diasporic community formation. It also hinges on ethical concerns about data research of this kind, as it was in large part the experience of looking at and reflecting on these photographic images that heightened the research team's ethical sensitivities on this issue. Ultimately, although these images were meant to address a wider public, we decided not the include the photographs in this article. We elaborate on our other ethical choices in the next

section, presenting a broader reflection on the implications of conducting research with digital methods on a sensitive topic in which human subjects have a direct stake.

Figure 6: ImagePlots (small multiples) for images associated with the source query [badecakir] on Instagram and ['Bade Cakir'] on Google Images
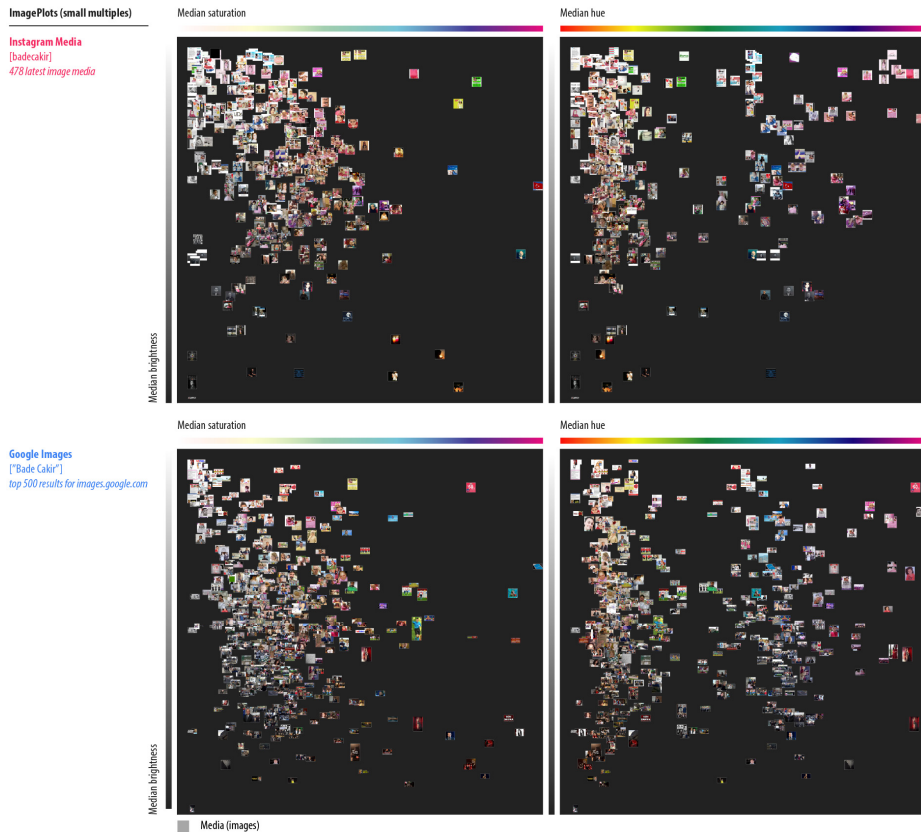
This 'direct visualization' (Manovich 2011) plots images as visual artefacts along two axes, thus rendering visual clusters of images with similar formalistic features (for example median brightness, median saturation and median hue). Data: 478 latest images on Instagram tagged with #badecakir on 20 April 2016 and top 500 image results for ['Bade Cakir'] on Google Images (images.google.com) on 11 February 2017. Data collection: DMI Instagram Scraper and DMI Google Image Scraper.

The social ethics of mapping

Data scraped and mined through the use of digital research tools increasingly raise ethical concerns in both academic and public debates, and the recent proliferation of articles, reports and guidelines on the topic (Clark et al. 2015; Metcalf et al. 2016; Zwitter 2014) testifies to contentions around the problem.[19] Starting at a fundamental level, we draw on the work of scholars who have raised questions about the ethical

implications of research inspired and informed by positivist ontologies, problematizing epistemological research frameworks that tend to understand data quantification as descriptive of reality and a vehicle of truth. We follow the critiques of big data as revolutionary (Mayer-Schönberger and Cukier 2013) and exposing it as validating a 'specific technological imaginary that rests on the mythological belief in the value of quantification and faith in its ability to model reality' (Dalton and Tatcher 2014). We position our work in relation to scholars who think about big data as a 'paradigm' (Beer 2016; Boellstorff and Maurer 2015; boyd and Crawford 2012), one that calls for a response in the form of critical digital media studies.

Hence, to avoid epistemological blindness towards the historical, social and technological embedding of big data 'discourse' (Beer 2016), we have proceeded from the assumption that knowledge is never merely descriptive or objective, but always processed through human intervention and interpretation. This has meant acknowledging the context-dependent character of digital data interpretation (Taylor 2015; van der Vlist 2016) and the role of power hierarchies in the realm of subjects' production, and researchers' processing, of such data. Nevertheless, relatively few of the scholars engaging in critical discussion of the big data ethics connect these oftentimes abstract reflections to practical methodological considerations for conducting research with digital tools that scrape data from public web platforms. In this concrete regard, we found feminist and ethnographic research approaches valuable for avoiding de-contextualized analyses of data, bringing up questions of voice (Gajjala 2013), temporalities, location and absence of data (Dourish et al. 2016), and trust, reflexivity and researcher's accountability towards participants (Leurs 2017; Pink et al. 2016; Preissle and Han 2012). Such perspectives informed our reflections on the relationship between researcher and researched in this case, and led us to think about ethical concerns over privacy in terms of data *ownership* (Metcalf and Crawford 2016). This perspective underpinned our decision to involve Bade's father, Muhammet Çakır, in the research process as a respondent, issue expert and consenter.

The decision to interview Muhammet and show him our digital data-based findings opened the way for introducing elements of a 'non-digital-centric' perspective on digital practices, based on the idea that to understand digital phenomena we needed to investigate the everyday, contextual worlds that sustain their use (Pink et al. 2016). As a visibly active and extremely invested actor in the digital and non-digital spread of this issue, we considered Muhammet uniquely positioned to (counter-)narrate the history of the data that our digital methods analysis had generated about Bade's issue. And while we do not claim that our choice ethically neutralizes our investigation, we are convinced that Muhammet was able to speak 'with the data, through the data and about the data' (Dourish et al. 2016), offering his lived experience in a way that enriched the analysis of the meaning-making process around the issue. This participative aspect of our research was part of our way of acknowledging 'the collaborative ways in which knowledge is made in the ethnographic process' (Pink et al. 2016), and it is for all these reasons that Muhammet Çakır's voice and consent regarding this research was important for us.

Integrating digital methods with data ethnography

We conducted our semi-structured interview with Bade's father, Muhammet Çakır, in December 2016 in Rotterdam. As with the digital data, we analysed the interview in terms of our interest in Turkish–Dutch diasporic group formations. We noted groupings mentioned in relation to Turkish friends and family, Turkish migrant institutions in the Netherlands, ethnic minority-hood in the Netherlands, and a global Turkish population. Most notably, these mentions contained recurrent references to numerical data, ranging from epidemiological statistics, numbers from international stem cell databases of registered donors, actual donor numbers, and the audience reach of news outlets and social media platforms. This finding was not only relevant to our interest in the role of different data types in this key actor's experience of the issue's spread, but it also revealed how data mediate the way contemporary diasporas come into being for their members.

Despite our interest in media-related and digital user data, we found that the kind of numerical data mentioned most frequently was numbers of donors and donor registrations. If we add to this the instances in which medical databases and epidemiological statistics were mentioned, these references account for most of the mentions of numerical data in the interview, despite the interviewers not having elicited responses about these medical data categories. Hence, the most meaningful way in which this key issue actor engaged with and experienced quantitative data in the context of this issue's overall development was through medical data, namely the data with the closest relation possible to his goal of saving Bade's life. All other kinds of numerical data – on the reach of social media platforms and news media outlets – were given meaning because they could potentially increase the donor pool and chances of a match for Bade and other patients in the Netherlands with a Turkish background. This issue's intensity for Muhammet and many around him brought the Turkish diaspora into being in a specific way around purpose-oriented engagement with specific kinds of data. And this context puts into perspective the role of web data in people's wider lived realities.

Relationships to close family and access to extended networks in media outlets are another non-digital facet of how the issue spread online. For instance, a friend of Muhammet's passed the item to his contacts in Turkey, after which the issue entered the Turkish-language news, both in the diaspora and subsequently in various Turkish-based news publications. It reached '80,000 people in the first five days', he said about a photo that Dutch-based Turkish journalists took of Bade in the hospital. He even mentioned asking some of the Dutch publications that interviewed him about their Turkish–Dutch audience demographics. This interaction between close social contacts and audience data shaped how Muhammet continued his social media work for the Bade foundation after his daughter's death, in turn also reflecting his understandings of how the community is mobilized, as he stated:

> Sometimes, I write something and my mother or my sister says: this is what you should put online! People still live and feel Bade. If you put something about Bade, then actually 50,000 or 60,000 people read it. If you put something about Rana [another Turkish girl with the same illness], only 20,000 people read it.

Apparently, Bade has more connection with people than Rana. … Bade gets more attention than other patients. For me, it doesn't help me anymore. But I see that telling something for Bade gets more attention than the new patients.

The interview data also revealed two significant temporal elements of how the issue's spread was experienced in addition to the timeline mapping (Figure 4a–b). First, Muhammet described the time pressure he experienced in mobilizing donor registrations by citing the tensions between the time taken for the labs to analyse results, the time Bade was predicted as having left to live, and the time that community mobilization took. Second, time was an important dimension of how Muhammet saw the changes in the engagement of those around him, describing how at the beginning – before the issue gained media reach – members of the community, organizations, and mosques were more reluctant to get involved. He specifically referred to the turnaround in one Turkish community organization's willingness eventually to organize a local donor drive event, one that ended up increasing the pool substantially. His frustration with this institutional reluctance extended to Dutch medical and governmental institutions such as the Dutch Ministry of Health, Welfare and Sport[20] (whose minister he described meeting) and the donor bank itself (whose processing speeds he raised concerns about). He described the news media attention as the cause for the turnaround in some of the local and national institutions' attention for the issue, for instance stating that:

> When people started to hear about the news, then there was attention coming from everywhere. We were asking wedding venues before to give us their places during the day. They didn't use the venues during the day. But they always said no. But, when people heard about it, everyone started to call. [They said] venues will be open for you, come and do your events. Then I started to choose the places. … When the events took place, there had to be one of us to welcome people and to help them with registration on computers. Friends from Türk Holland Gönüllüler Vakfı helped a lot.

For Muhammet this important moment signalled a peak in registration, which was the product of practical facilitation of the activity of communally registering online. And while we have mapped the aggregate temporal peaks of engagement with the issue on social media, the feelings of being supported by, proud of, or frustrated and angry with the wider Turkish community are an emotional dimension of this issue's spread that bore on Muhammet directly and could not be quantified in the data. Muhammet's explicit mentions of 'mixed emotions' were also linked to questions of control (and lack thereof) over the online spread of visual content. Through the strategic use of media, Muhammet effectively maintained this control over the number and kind of videos and photographs displayed and circulated in news and social media, explicitly wanting to avoid becoming media fodder or sensationalized, as we see in his description of the high-profile, televised visit that a Turkish Prime Minister's wife paid to Bade:

> The ambassador heard about Bade so they wanted to put it [a visit] together. They called me, but I said if it's only to fill her gap, no. But if she wants to come here and explain the situation. … If she mentions it and thousands of people register, then it's okay.

Conversely, losing control of the media circulation of the issue caused Muhammet emotional distress, reflecting how the mobilization caused ownership of visual data shared online to shift subtly from Muhammet to his audience, giving him a sense of discomfort and powerlessness. As he stated: 'those pictures [of Bade, ill in the hospital] are moving through the internet and I cannot take them back. It feels so strange.' While enriching the visualized data with additional dimensions, the interview constitutes part of our alternative digital mapping process and, interestingly, reflects no distinct division between 'online' and 'offline' data or media. Instead, the narrative reflects the emergence of different issue spheres in a way that resonates with the earlier sections of our analysis. A news-related sphere across different journalistic media outlets seems to emerge separately, while the two main social media platforms each reflect a different spread of and engagement with the issue. This makes it interesting that Muhammet spoke about Facebook as the sphere in which he was active, and Instagram as being predominantly the domain of his wife,[21] suggesting that how the issue engaged gendered patterns of use that revealed platform affordances may indeed have shaped some of the differences in how the issue spread on different platforms. So far, the anthropological work on social media has largely highlighted the importance of understanding social and cultural patterns that shape social media use in everyday contexts. For instance, Costa's (2016) ethnographic study in Southeast Turkey reveals how users fit social media usage practices into already existing social and cultural forms. In this case, Muhammet's narrations of his own and his wife's roles in relation to the different media platforms through which the issue was spread suggested the gendered sociality of engagement with this significant issue translating to social media use.

Conclusion

We have argued for the development of empirical 'mattering maps' as an innovative mode of mapping diaspora digitally, one that draws on ethnographic and feminist research principles to reflect the temporal and locational contingency of diasporic formation, while also acknowledging the partiality and situatedness of the mapping exercise itself. Since one cannot comprehensively, objectively, or definitively map engagement around issues that mobilize diaspora with any digital tool, one needs to trace how a diaspora comes into mediated being through specific engagements with data and uses of digital media within lived daily existence. At its core, this mixed methods approach puts into practice the assertion about 'big data' studies by Boellstorff and Maurer (2015) that 'addressing researcher subjectivity makes research more scientific, robust, and ethical'. We have developed our specific model of mixing methods by mapping a community's engagement with an issue through an empirical case study embedded within a conceptual discussion of diasporic formation processes, but one can potentially apply

this model of integrating digital methods to ethnographic research in other research contexts where digital media use around an issue is studied as embedded within everyday social and media practices.

The distinct issue spheres that emerged through our analysis helped us to organize our findings and showed the importance of tracing the issue across platforms and of identifying how digital diasporic practices are not only temporally and spatially produced but also produced in accordance with specific platforms and the kinds of usage they afford. At the same time, the interview data helped us to make the case for tracing the issue into the realm of lived experience, stressing not only how 'the affordances of a particular technology shape what people do' but also 'how people, through their social practices, shape what the technology can do' (Hine 2016). This is reflected in the integration of how the online and offline are thought about and put into practice, but perhaps more importantly it allowed us to show how engagement with different forms of quantifiable data and intense emotional experiences are both equally and inherently part of people's contemporary lives. There is therefore an urgent need to develop further research frameworks for understanding affective/emotional dimensions of (audio) visual digital content, how users who are also inundated with various forms of data experience and interpret them, and how these factors together come to mediate the formation of diasporas in a digital age.

While the recent ethnographic scholarship on digital media and the social world has begun to integrate quantitative data into ethnographic analysis (see Hjorth et al. 2016; Miller and Sinanan 2017; Miller et al. 2016) our case study analysis has contributed to the growing area of mixed-methods research by demonstrating concrete, methodological possibilities and necessities for integrating the use of *digital methods tools* programmed for large-scale data collection with ethnographic research and ethical practice. By pointing out these possibilities, we aim to demonstrate that this alternative mapping exercise could be further embedded within more traditional, in-depth ethnographic investigation, in the service of building theory that engages with longer-running debates on diaspora from anthropological and cultural studies perspectives.

Acknowledgements

Notes

1. Google Search is the market leader for all three result sets (Return on Now 2015).
2. https://www.matchis.nl/word-stamceldonor/patient-en-donorverhalen/kleine-dappere-bade.
3. Dutch: Turkse Vrijwilligers Stichting.
4. https://www.facebook.com/752828048195495/.
5. English: 'Marrow for Bade – a favour for bade'.

6. https://www.facebook.com/1546625252320085/.
7. https://www.facebook.com/176725159173895/.
8. https://tools.digitalmethods.net/netvizz/instagram/.
9. English: 'stem cell for Bade'.
10. English: 'Bade became an angel'.
11. English: 'stem cell'.
12. English: 'the little princess'.
13. Our interview with Bade's father later revealed that a family member did not own the account.
14. Total number of posts associated with '#badecakir' per month: January 2016: 97 posts; February 2016: 214 posts; March 2016: 63 posts.
15. This number is obtained through a keyword search on 20 April 2016 via the site's front-end after the API was closed.
16. http://badefoundation.com/.
17. http://sonhaber.eu/yasam/bade-iki-hayat-kurtardi/.
18. 'Repost for Instagram' is a mobile application that allows users to repost images shared by other users on Instagram. As Figure 3b shows, '#repost' was one of the most circulated hashtags as the application automatically tags media with that hashtag.
19. Recent examples of unethical appropriation of digital data that have sparked widespread debate include Facebook's emotional contagion research project in which scholars manipulated content shown on the News Feed of 689,003 unaware Facebook users (Kramer et al. 2014).
20. Dutch: Ministerie van Volksgezondheid, Welzijn en Sport.
21. Research on users of different social media platforms shows how Facebook is a platform with a balanced gender usage while women are more likely to use Instagram (Duggan et al. 2015). There are also studies that analyse specific gendered uses of Instagram on issues like pregnancy (Tiidenberg 2015; Tiidenberg and Baym 2017), and motherhood (see Lupton et al. 2016). In examining the online virality of Aylan Kurdi's iconic image, Farida Vis (2015) mentions the work of Rose (2010) on family photography to counterpoise the mother's absence from photographic records with her role in creating those records. In this sense, it is interesting to think about Instagram's specific photographic affordances and their relationship with gendered usage of photography.

References

Anthias, F. (1998) 'Evaluating "diaspora": beyond ethnicity?', *Sociology*, 32 (3), 557–80, doi: 10.1177/0038038598032003009.

Beer, D. (2016) 'How should we do the history of big Data?', *Big Data & Society*, 3 (1), 1–10, doi: 10.1177/2053951716646135.

Boellstorff, T. and B. Maurer (eds) (2015) *Data, now bigger and better!* Chicago, IL: Prickly Paradigm Press.

boyd, d. and K. Crawford (2012) 'Critical questions for big data', *Information, Communication & Society*, 15 (5), 662–79, doi: 10.1080/1369118X.2012.678878.

Brah, A. (1996) *Cartographies of diaspora: contesting identities*, New York: Routledge.

Brinkerhoff, J. M. (2009) *Digital diasporas: identity and transnational engagement*, Cambridge: Cambridge University Press.

Clark, K., M. Duckham, M. Guillemin, A. Hunter, J. McVernon, C. O'Keefe, C. Pitkin, S. Prawer, R. Sinnott, D. Warr and J. Waycott (2015) *Guidelines for the ethical use of digital data in human research*, Melbourne: University of Melbourne, available at: www.carlton connect.com.au/wp-content/uploads/2015/06/Ethical-Use-of-Digital-Data.pdf.

Clifford, J. (1994) 'Diasporas', *Cultural Anthropology*, 9 (3), 302–38, doi: 10.1525/can.19 94.9.3.02a00040.

Costa, E. (2016) *Social media in southeast Turkey: love, kinship and politics*, London: UCL Press.

Dalton, C. and J. Thatcher (2014) 'What does a critical data studies look like, and why do we care?', *Society & Space*, available at: http://societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/.

Diminescu, D., M. Bourgeois, M. Renault and M. Jacomy (2011) 'Digital Diasporas Atlas: exploration and cartography of diasporas in digital networks', in *Proceedings of the fifth international AAAI conference on weblogs and social media (ICWSM-11)*, Palo Alto, CA: Association for the Advancement of Artificial Intelligence, 657–8, available at: https://pdfs.semanticscholar.org/66c7/79a277fb66f255378476d5ee97a4a045eb14.pdf.

Diminescu, D. and B. Loveluck (2015) 'Traces of dispersion: online media and diasporic identities', *Crossings: Journal of Migration Culture*, 5 (1), 23–39, doi: 10.1386/cjmc.5.1.23.

Dourish, P., E. Gómez Cruz, H. Horst, D. Lupton, S. Pink, J. Postill, S, Sumartojo and D. Verhoeven (2016) 'Data Ethnographies 4: data stories', available at: https://dataethnographies.com/paper-iv-data-stories/.

Duggan, M., N. B. Ellison, C. Lampe, A. Lenhart and M. Madden (2015) 'Social media update 201.' Pew Research Center webpage, available at: www.pewinternet.org/2015/01/09/social-media-update-2014/.

Fuchs, C. (2017) 'From digital positivism and administrative big data analytics towards critical digital and social media research!', *European Journal of Communication*, 32 (1), 37–49, doi: 10.1177/0267323116682804.

Gajjala, R. (ed.) (2013) *Cyberculture and the subaltern: weavings of the virtual and real*, Lanham, MD: Lexington Books.

Georgiou, M. (2006) *Diaspora, identity and the media: diasporic transnationalism and mediated spatialities*, New York: Hampton Press.

Gilroy, P. (1993) *The black Atlantic: modernity and double consciousness*, Cambridge, MA: Harvard University Press.

Gomes, C. (2016) *Transient mobility and middle class identity: media and migration in Australia and Singapore*, Singapore: Palgrave Macmillan.

Gonzalez-Polledo, E. and J. Tarr (2016) 'The thing about pain: the remaking of illness narratives in chronic pain expressions on social media', *New Media and Society*, 18 (8), 1455–72, doi: 10.1177/1461444814560126.

Grossberg, L. (2010) 'Affect's future: rediscovering the virtual in the actual', in M. Gregg and G. J. Seigworth (eds) *The affect theory reader*, Durham, MD: Duke University Press, 309–38.

Hall, S. (2003) 'Cultural identity and diaspora', in J. E. Braziel and A. Mannur (eds) *Theorizing diaspora: a reader*, Malden: Blackwell Publishing, 233–46.

Hine, C. (2015) *Ethnography for the internet: embedded, embodied and everyday*, London: Bloomsbury Academic.

Hine, C. (2016) 'From virtual ethnography to the embedded, embodied, everyday internet', in L. Hjorth, H. A. Horst, A. Galloway and G. Bell (eds) *The Routledge companion to digital ethnography*, New York: Routledge, 21–8.

Hjorth, L., H. A. Horst, A. Galloway and G. Bell (eds) *The Routledge companion to digital ethnography*, New York: Routledge.

Johnson, M. and D. McKay (2011) 'Mediated diasporas: material translations of the Philippines in a globalized world', *Southeast Asian Research*, 19 (2), 181–96, doi: 10.5367/sear.2011.0047.

Kok, S. and R. Rogers (2016) 'Rethinking migration in the digital age: transglocalization and the Somali diaspora', *Global Networks*, 17 (1), 23–46, doi: 10.1111/glob.12127.

Kozinets, R. V. (2009) *Netnography: doing ethnographic research online*, London: SAGE.

Kramer, A. D. I., J. E. Guillory and J. T. Hancock (2014) 'Experimental evidence of massive-scale emotional contagion through social networks', *PNAS*, 111 (24), 8788–90, doi: 10.1073/pnas.1320040111.

Leurs, K. (2017) 'Feminist data studies: using digital methods for ethical, reflexive and situated socio-cultural research', *Feminist Review*, 115 (1), 130–54, doi: 10.1057/s41305-017-0043-1.

Lupton, D., S. Pedersen and G. M. Thomas (2016) 'Parenting and digital media: from the early web to contemporary digital society', *Sociology Compass*, 10 (8), 730–43, doi: 10.1111/soc4.12398.

McKay, D. (2012) *Global Filipinos: migrants' lives in the virtual village*, Bloomington: Indiana University Press.

Madianou, M. and D. Miller (2012) *Migration and new media: transnational families and polymedia*, London: Routledge.

Manovich, L. (2011) 'What is visualisation?', *Visual Studies*, 26 (1), 36–49, doi: 10.1080/1472586X.2011.548488.

Marres, N. (2015) 'Why map issues? On controversy analysis as a digital method', *Science, Technology and Human Values*, 40 (5), 655–86, doi: 10.1177/0162243915574602.

Mayer-Schönberger, V. and K. Cukier (2013) *Big data: a revolution that will transform how we live, work, and think*, New York: Houghton Mifflin Harcourt.

Metcalf, J. and K. Crawford (2016) 'Where are human subjects in big data research? The emerging ethics divide', *Big Data & Society*, 3 (1), 1–14, doi: 10.1177/2053951716650211.

Metcalf, J., E. F. Keller and d. boyd (2016) *Perspectives on big data, ethics, and society*, New York: The Council for Big Data, Ethics, and Society.

Miller, D. and J. Sinanan (2017) *Visualising Facebook: a comparative perspective*, London: UCL Press.

Miller, D., E. Costa, N. Haynes, T. McDonald, R. Nicolescu, J. Sinanan, J. Spyer, S. Venkatraman and X. Wang (2016) *How the world changed social media*, London: UCL Press.

Nakamura L. and P. Chow-White (eds) (2012) *Race after the Internet*, New York: Routledge.

Pink, S., H. Horst, J. Postill, L. Hjorth, T. Lewis and J. Tacchi (2016) *Digital ethnography: principles and practice*, London: Sage.

Preissle, J. and Y. Han (2012) 'Feminist research ethics', in S. N. Hesse-Biber (ed.) *The handbook of feminist research: theory and praxis*, second edition, Thousand Oaks, CA: Sage, 583–605.

Return on Now (2015) '2015 search engine market share by country', Return on Now (internet marketing consulting service) webpage, available at: http://returnonnow.com/internet-marketing-resources/2015-search-engine-market-share-by-country/.

Rieder, B. (2013) 'Studying Facebook via data extraction: the Netvizz application', in *WebSci' 13: Proceedings of the 5th Annual ACM Web Science Conference*, New York: ACM Press, 346–55, doi: 10.1145/2464464.2464475.

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

Rogers, R., N. Sánchez-Querubín and A. Kil (2015) *Issue mapping for an ageing Europe*, Amsterdam: Amsterdam University Press.

Rose, G. (2010) *Doing family photography: the domestic, the public and the politics of sentiment*, Farnham: Ashgate.

Taylor, L. (2015) 'Towards a contextual and inclusive data studies: a response to Dalton and Thatcher', *Society & Space*, available at: http://societyandspace.org/2015/05/12/towards-a-contextual-and-inclusive-data-studies-linnet-taylor/.

Tiidenberg, K. (2015) 'Odes to heteronormativity: presentations of femininity in Russian-speaking pregnant women's Instagram accounts', *International Journal of Communication*, 9 (13), 1746–58, doi: 1932–8036/2015FEA0002.

Tiidenberg, K. and N. K. Baym (2017) 'Learn it, buy it, work it: intensive pregnancy on Instagram', *Social Media + Society*, 3 (1), 1–13, doi: 10.1177/2056305116685108.

Trandafoiu, R. (2013) *Diaspora online: identity politics and Romanian migrants*, New York: Berghahn Books.

van der Vlist, F. N. (2016) 'Accounting for the social: investigating commensuration and big data practices at Facebook', *Big Data & Society*, 3 (1), 1–16, doi: 10.1177/2053951716631365.

Vis, F. (2015) 'Examining the hundred most shared images of Aylan Kurdi on Twitter', in F. Vis and O. Goriunova (eds) *The iconic image on social media: a rapid response to the death of Aylan Kurdi*, Sheffield: Visual Social Media Lab, 27–30.

Weltevrede, E. J. T. (2016) 'Repurposing digital methods: the research affordances of platforms and engines', Ph.D. thesis, University of Amsterdam, available at: http://hdl.handle.net/11245/1.505660.

Wimmer, A. and N. Glick Schiller (2002) 'Methodological nationalism and beyond: nation-state building, migration and the social sciences', *Global Networks*, 2 (4), 301–34, doi: 10.1111/1471-0374.00043.

Zwitter, A. (2014) 'Big data ethics', *Big Data & Society*, 1 (2), 1–6, doi: 10.1177/2053951714559253.