CHAPTER 5

Big Data-Driven Health Surveillance

The emergence of research using big data for public health surveillance is directly related to the vast, diverse data generated by individuals online. On Twitter, many users publicly post about their medical conditions, medication and habits related to self-care. By 'liking' content, Facebook users indicate their eating habits or physical (in-)activity. It is common to search the internet for information on experienced or observed diseases and symptoms. Some users sign up for online communities to exchange their personal knowledge of and struggles with illness, and some even track their physical movements and physiological signals with wearable fitness devices. Such data have come to play a role in research on public health surveillance.

When drawing on such data, especially when applying for funding and when publishing results, researchers articulate ethical arguments and validity claims contending the normative rightness of their approaches. Some of these claims will be examined in the following chapter, with specific regards to research on big data-driven public health surveillance. Important trends in this field are approaches monitoring social media, search behaviour and information access. As an alternative to mining data without users' consent, possibilities of health prosumption and participatory epidemiology are being explored.

Social media monitoring as contribution to public health surveillance. On social networking sites such as Facebook or microblogging platforms like Twitter, users post and interact with potentially health-relevant information. They may, for example, casually post about their health conditions or indicate interests and (e.g. dietary or sexual) habits which may be health-related. This sharing of information facilitates research drawing on social media data collected by tech corporations. Such research may be conducted by scientists employed at universities and (inter-)governmental institutes, and potentially in collaboration with employees of tech corporations.

Search behaviour and information access. Due to their widespread use, search engines (most notably Google) act as main gateways to online information. Among many other things, users enter queries which may be health related and potentially allow for insights into their health conditions as well as experiences concerning, for example drugs, treatments, health providers, or physicians. Such search queries, however, are not only entered on websites which are predominantly search services. Users may also search for persons and access content related to their interests on social networking sites. Therefore, these kinds of data also play a role for the first category mentioned above. Data emerging from users' search queries can have high biomedical value in various regards. Therefore, they have been used as means for public health monitoring. Such datasets have only rarely been provided as open data, since early attempts demonstrated the difficulties of anonymisation (Zimmer 2010; Arrington 2006). Related studies have been mainly conducted by scientists employed at tech corporations, or in a few cases in public-private collaboration.

Health prosumption and participatory epidemiology. Social networking sites allow for and encourage users' participation; for example, in the form of content contributions or communal support. These forms of 'prosumption' have also facilitated the development of health platforms that engage users in ways leading to biomedical big data. In this context, research and projects have emerged which aim at developing platforms or applications needed to collect data. They are meant to create possibilities for individuals' deliberate involvement in public health surveillance as a form of 'participatory epidemiology' (Freifeld et al. 2010). Such initiatives emerged in university contexts, as part of (inter-)governmental institutions and/or businesses.

In the following subchapters, I will mainly investigate cases of social media monitoring and big data use in research on public health surveillance. I will highlight three domains: first, data retrieved from users who provide indications of physical/health conditions and behaviour, voluntarily or involuntarily, knowingly or unknowingly; secondly, data retrieved from users' interaction with social media content and features; thirdly, data retrieved, combined, and mapped based on multiple online sources. I will refer to the relevance of search queries as a data source, as well as to examples of 'participatory epidemiology'. The latter will be described in less detail though, since related approaches do not necessarily classify as big data.

High-Risk Tweets: Exposing Illness and Risk Behaviour

Especially early on, efforts in digital disease detection focused on the surveillance of influenza (e.g. Eysenbach 2006; Polgreen et al. 2008; Ginsberg et al. 2008; Signorini et al. 2011). The topical focus on influenza or influenza-like-illness (ILI) owes partly to to its widespread occurrence, but influenza is also

an illness that sufferers/users tend to be comparatively open about discussing. A person who states to suffer 'from the flu' on social networking sites is relatively likely to experience sympathy (possibly also disbelief or disinterest). Individuals posting about suffering from symptoms related to their infection with the human immunodeficiency virus (HIV) may instead be subjected to stigma and discrimination.

Certain infectious diseases, such as HIV/AIDS, are known to be highly stigmatising for affected patients (Deacon 2006). This also applies to mental illnesses such as schizophrenia (Crisp et al. 2000). Affected individuals are less likely to openly and lightly post explicit information on their health condition in cases of highly stigmatised conditions. This also has implications for the accessibility of information and data regarding these diseases. It implies that certain disease indicators are reflected only implicitly and not explicitly in users' content. Despite these complicating conditions regarding big data on diseases such as HIV, studies have examined how social media can be used to monitor relevant factors. In comparison to research on big data relevant to influenza monitoring, in these cases the focus is less on articulations of symptoms, but on content indicative for risk behaviour. A difference concerning the data sources is therefore that an individual posting about or searching for information on flu symptoms is more likely to be aware what this content signifies. In comparison, a person posting about certain habits which can be classified as, for example, drug- or sex-related risk behaviour is perhaps unaware that these posts may be indicators of health risks.

As part of the BD2K funding scheme 'Targeted Software Development', several research projects explore how social networking sites could play a role in countering infectious diseases. Broadly speaking, they examine how online data may reflect users' health behaviour and conditions. Examples for projects active in 2017/18 are 'Mining the social web to monitor public health and HIV risk behaviors' (Wang et al. n.d.)[58] and 'Mining real-time social media big data to monitor HIV: Development and ethical issues' (Young et al. n.d.)[59]. Also, outside of the BD2K scheme, funding has been granted to projects such as 'Online media and structural influences on new HIV/STI Cases in the US' (Albarracin et al. n.d.)[60]. The responsible interdisciplinary research teams consist of epidemiologists, computer and data scientists, public health researchers and psychologists. Similar projects have been launched with regards to mental illness monitoring, for example 'Utilizing social media as a resource for mental health surveillance' (Conway n.d.)[61]. The analysis below will, however, focus on social media monitoring of content considered relevant for HIV/AIDS risk factors.[62]

Research in this field has as yet received little public attention, possibly due to the fact that it has emerged relatively recently. Moreover, it could be speculated that these research practices were not found to be controversial or problematic by journalists or other observers. In any case, insights have so far mainly been communicated via academic outlets, and targeted at researchers or

public health professionals/institutions. Therefore, the arguments brought forward in this context are likewise predominantly established by researchers and not by external observers such as journalists or private individuals. Drawing on Habermas' notion of validity claims, especially with regards to 'normative rightness', but also 'truth' and 'authenticity', the following sections elaborate on the ethical arguments raised in big data-driven approaches to monitoring of HIV/AIDS risk behaviour.

HIV/AIDS risk behaviour refers, for example, to drug consumption which can be hazardous to health, such as the sharing of needles or unprotected sex. To examine how such factors could be monitored via social networking sites, all the projects mentioned above make use of Twitter data. As described in Chapter 3, the microblogging platform broadly allows for open data access. Building on Twitter data, Wang et al. (n.d.) '[…] propose to create a single automated platform that collects social media (Twitter) data; identifies, codes, and labels tweets that suggest HIV risk behaviors'. The platform is meant to be used as tool and service by stakeholders such as HIV researchers, public health workers and policymakers.

The project starts from the hypothesis that certain tweets indicate that individuals intend to or did engage in sex- and drug-related risk behaviour. Some of those tweets can be (roughly) geographically located and enable the monitoring of certain populations (see Young, Rivers, and Lewis 2014). The significance of retrieved data is assessed by combining them with data from established public health surveillance systems as provided by, among others, the US Centres for Disease Control and Prevention (CDC) or the WHO. Wang et al.'s project is particularly focused on automating the processes leading to an identification of potentially relevant data.

In a related paper, the involved scholars acknowledge the importance of preventing their research being linked back to individual persons, since this could lead to stigmatisation (Young, Yu and Wang 2017: 130). For this reason, only a partial list of keywords significant as risk factor indicators has been provided. While stating that '[a] large and growing area of research will be focused on how to address the logistical and ethical issues associated with social data' (130), the authors do not address those issues in detail themselves. However, the project by Young et al. (n.d.; as mentioned before, the scientist was also involved in the study mentioned above) refers explicitly to the relevance of ethical concerns. Methodologically, it moves beyond an exploration of technical challenges. It adds qualitative interviews with '[…] staff at local and regional HIV organization and participants affected by HIV to gain their perspectives on the ethical issues associated with this approach' (Young et al. n.d.).

The two projects highlight typical, insightful approaches to ethical issues in big data research. Concerns regarding the normative rightness and risks of big data-driven studies are framed as challenges to be overcome in future research; they are, however, not seen as reasons to explore beforehand which moral issues may arise. This innovation-driven approach also reflects the conditions

under which biomedical and life scientists compete for funding. In the above-mentioned cases, it remains to be clarified if and how such research may affect social media users, for example by becoming accused of or associated with presumed HIV/AIDS risk behaviour. But, practically speaking, flagging severe ethical issues may undermine the perceived feasibility and 'fundability' of a research project.

Moreover, an emphasis on ethical questions appears less likely to receive funding in schemes explicitly targeted at software development. At the same time, these dynamics seem related to a lack of ethical guidelines concerning biomedical big data, commonly ensured by institutional/ethical review boards (I/ERB). Ethical decision-making processes for big data-driven public health research operate currently according to negotiated rationales, such as necessity versus the obsolescence of informed consent (see Chapter 3). This also puts involved researchers at risk of public, morally motivated scandalisation and distrust.

Already in traditional *Infectious Disease Ethics* (IRD), a sub-discipline of bioethics concerned with ethical issues regarding infectious diseases, Selgelid et al. (2011) observed comparable tensions between scientists and philosophers, particularly ethicists. While scientists experienced certain moral expectations as unrealistic and oblivious of research realities, philosophers perceived scientists' consideration of ethical issues as naïve. This in turn was countered by scientists with the objection '[…] we are not ethicists, we're just describing an ethical issue we have observed' (Selgelid et al. 2011: 3).

A view of ethics as an 'ex post' perspective is thus not a feature characteristic for big data-driven research, but rather a tendency which can be found in novel, emerging research fields. Moreover, it brings forward the normative claim that ethics cannot be demanded as key, analytic expertise from (data) scientists. Such dynamics have facilitated a 'pacing problem' in innovative research and a '[…] gap between emerging technologies and legal-ethical oversight' (Marchant, Allenby and Herkert 2011). In fast-changing technological cultures, ethical debates often lag behind (see also Wilsdon and Willis 2004). This point hints not only at the importance of strengthened collaboration and mediation between ethicists and scientists, but also at the need for research skills relevant to projects' ethical decision making and increased public outreach.

A recurring ethical, contested issue in this context, as already indicated in Chapter 3, is the question of informed consent. While Young et al. (n.d.) deliberately incorporate stakeholders such as public health professionals and individuals affected by HIV, the role of other users creating data receives little consideration. It has been pointed out that posting content on social media does not necessarily correspond with users' awareness of possible, future uses. Furthermore, users often have little means of privacy management once they opt-in for using certain platforms (Baruh and Popescu 2015; Antheunis, Tates, and Nieboer 2013; boyd and Ellison 2007). Research drawing on such data affects users as it claims access to personal data whose use has not been

explicitly authorised by the respective users. This has implications for the societal appreciation of personal autonomy.

The tendency to portray informed consent as neglectable is linked to the common framing of big data approaches as 'unobtrusive', i.e. occurring seemingly without intervening with individuals' activities (see also Zwitter 2014). For example, the scientists involved in the project 'Online media and structural influences on new HIV/STI Cases in the US' (Albarracin et al. n.d.) examined tweets as possible indicators of HIV prevalence in (2079 selected) US counties. Similar to the projects by the PIs Wang and Young, Albarracin et al. also focus on potential links between linguistic expressions on Twitter and HIV prevalence in a population. The authors describe their retrieval of 150 million tweets, posted between June 2009 and March 2010, as '[…] an unobtrusive, naturalistic means of predicting HIV outbreaks and understanding the behavioral and psychological factors that increase communities' risk' (Ireland et al. 2015). In this context, 'unobtrusive' is used in the sense that the data collection does not interfere with users' social media practices.

Implicitly, this interpretation of unobtrusiveness is used as a claim to normative rightness. The normative assumption brought forward in this context is that an approach may be considered unobtrusive because the involved subjects are not necessarily aware that their data are being collected. This claim to the normative rightness and preferability of such approaches is paired with the argument that it produces 'undistorted' and 'better' data, a validity claim to truth. Considering that the latter argument has been challenged as a discursive element of a 'digital positivism' (Mosco 2015) and 'dataism' (van Dijk 2014), these validity claims to normative rightness and truth alike are questionable. Ethically, it implies a misleading understanding of (un-)obtrusiveness which is then presented as advantageous. Methodologically, its claims to reduce distortion appear questionable in the light of research on algorithmic bias (see Chapter 3).

These entanglements between claims to normative rightness and truth are decisive. With regards to *Infectious Disease Ethics*, Selgelid et al. (2011) state that commonly '[r]estrictions of liberty and incursions of privacy and confidentiality may be necessary to promote the public good' (2). But implied measures such as quarantine and mandatory vaccinations usually apply to 'extreme circumstances' (2) or consequences. Moreover, in assessing whether certain ends justify the means, the approaches' effectiveness becomes an important concern. Claims for the normative rightness of social media monitoring for public health surveillance therefore also need to be assessed in light of their claims to effectiveness.

As discussed in Chapters 2 and 3, valid concerns have been raised regarding factors biasing and distorting big data. In the case of the abovementioned studies, two aspects especially should be considered: first, the alterability of corporate big data economies; and second, the fluidity of user behaviour. Both aspects translate into matters of sustainability, reliability, and accuracy. While

prominent figures in the field of health informatics such as Taha A. Kass-Hout[63] have declared that "Social media is here to stay and we have to take advantage of it,' […]' (Rowland 2012), neither the platforms nor the corporations owning them are static. Even though Twitter has survived prognoses for its bankruptcy made in 2016 (Giannetto 2015) and it has been said that 'Twitter Inc. can survive' (Niu 2017), the company is struggling to achieve profitability (Volz and Mukherjee 2016).

While one may oppose the possibility that Twitter may be discontinued, given its popularity, it is certainly likely that its data usage conditions will continue to change. This has already occurred in the past, as pointed out by Burgess and Bruns (2012) and Van Dijck (2011). Amendments in Twitter's APIs, making certain data inaccessible, imply that research projects relying on the microblogging platform as their main data source could not proceed as planned. This risk is especially significant when it comes to collaboration with start-ups, as demonstrated by other cases. For example, in February 2016, the Indiana University School of Nursing announced its collaboration with ChaCha, a question and answer online service ('IU School of Nursing and ChaCha partner' 2015).

The platform was available as a website and app. Users could ask questions which were then answered by guides, paid by the company on a contractor basis. It was launched in 2006, received an estimated $43-58 million venture capital within three years (Wouters 2009), first filed bankruptcy in 2013 (ChaChaEnterprises, LLC 2013), and ceased to exist in 2016 (Council 2016). In 2015 the company established a data sharing agreement with the Social Network Health Research Lab (Indiana University, School of Nursing). The researchers received a large (unspecified) dataset of user questions submitted between 2008 and 2012. The aim is/was to analyse questions pertinent to health and wellness, and to explore their implications for public health monitoring. While this one-off data donation still allows researchers to examine the material, follow-up studies involving more recent data would be impossible.

With regards to Twitter and other social networking platforms such as Facebook it has been frequently assumed and argued that privacy is not an ethical issue, because '[…] the data is already public' (Zimmer 2010, 313). In a critical paper on the use of Facebook data for research, Zimmer investigates the unsuccessful anonymisation of a data set and reveals 'the fragility of the presumed privacy of the subjects under study' (314). In a later article, Zimmer and Proferes (2014) oppose the dominant argument that users '[…] have minimal expectations of privacy (Crovitz, 2011), and as a result, deserve little consideration in terms of possible privacy harms (Fitzpatrick, 2012)' (170). When using Twitter, users can choose between either making all their tweets public or restricting access to authorised users. Tweets which are posted publicly are fed into Twitter's partly open data and can be accessed via API. The company itself has access to all tweets, published publicly or privately, as well as metadata, i.e. hashtags, page views, links clicked, geolocation, searches, and links between users (172). Zimmer[64] and Proferes (2014) show that despite Twitter's

seemingly straightforward, binary mechanism of public and private tweets, the platform's marketing generally evokes promises of 'ephemeral content sharing'.

As part of the Council for Big Data, Ethics, and Society,[65] established in 2014 as an initiative providing critical social and cultural perspectives on big data, a report by Uršič (2016) shows that in cases where civic users delete tweets or content, this material often remains part of retrieved datasets (5ff.). Coming back to the use of Twitter data for monitoring HIV/AIDS risk factors, the wish to delete personal tweets may occur especially once it transpires how certain content may be interpreted. One should also take into account that not only a platform's appearance, usage conditions and possibilities may be fluid, but that the same goes for users' behaviour. Once aware of the possibility that certain communications (even if only vaguely related to one's sex life, drug consumption, or social drive) may be interpreted as risk behaviour, this could alter users' content production.

Such a development is easily conceivable, given common prejudices towards and the stigmatisation of individuals' suffering from HIV/AIDS. And even without such an explicit intention to adjust behaviour to avoid discrimination, or the impossibility to find an insurer, individuals' interests and practices change. This means that content which might have implied drug- or sex-related risk behaviour may in the foreseeable future take on a different meaning. At this point, it is insightful to remember 'lessons learned' from the discontinuation of Google Flu Trends. In an article on 'big data hubris', Lazer et al. (2014) warn that the constant re-engineering of platforms such as Twitter and Facebook also means that '[…] whether studies conducted even a year ago on data collected from these platforms can be replicated in later or earlier periods is an open question' (1204). In addition, the authors stress the role of so-called 'red team dynamics' resulting from users' attempts to '[…] manipulate the data generating process to meet their own goals, such as economic or political gain. Twitter polling is a clear example of these tactics' (1204).

Comparable dynamics may not only occur due to activities aimed at deliberate manipulation, but also in cases where users react to current events or trends. As early as 2003, Eysenbach (see also 2006) underlined the possibility of 'epidemics of fear'. With this term, the author differentiates between digital data which may reflect that individuals are directly affected by a disease, and those that emerge because users may have heard or read about a health-relevant development. In the case of Google Flu Trends, for example, it is assumed that search queries indicating 'epidemics of fear' have acted as confounding factors, leading repeatedly to overestimations (Lazer et al. 2014, 1204): inter alia during the 2009 H1N1 pandemic (Butler 2013). For the abovementioned projects, aimed at employing social media monitoring as contributing to HIV/AIDS surveillance, this means that models developed based on research need to be constantly evaluated, adjusted, and recalibrated. One reason for this is that linguistic content which has been selected as a signifier of risk-behaviour may subsequently take on different meanings.

This applies to all projects drawing on social media data such as tweets which have been used for monitoring, for example, influenza or cholera,[66] but it seems notably relevant for projects that address stigmatised health conditions. Likewise, the politics behind the selection of certain content which is screened as being indicative for risk behaviour should also be considered. This relates particularly to emphasis on groups that are potentially 'high risk'. If we look for instance at concerns in a different area, regarding 'racial profiling' (Welsh 2007), it has been noted that discriminatory attention towards groups can foster selection and sampling bias. While this is not meant to query that HIV/AIDS research is especially relevant for certain vulnerable groups and individuals, the translation of this knowledge into linguistic criteria for big data-driven research may facilitate sampling biases in the chosen material.

With regards to observational epidemiology, Chiolero (2013) remarks that already in a 'pre-big data era' the trust in large-scale studies occasionally undermined methodological scrutiny. As the author observes, '[…] big size is not enough for credible epidemiology. Obsession with study power and precision may have blurred fundamental validity issues not solved by increasing sample size, for example, measurement error, selection bias, or residual confounding' (Chiolero 2013). Such methodological issues are, however, difficult or even impossible to assess for an external observer, since the ethical concerns regarding stigmatisation led to scientists' decision not to reveal linguistically significant keywords and data.

Similar variations of digital disease detection have also been used in response to natural disasters and humanitarian crises such as the 2010 Haiti earthquake and the subsequent cholera outbreak (Meier 2015; Chunara et al. 2012). On Twitter's tenth anniversary, UN Global Pulse praised the platform as '[…] one of the central data sources here at Global Pulse during our first years of implementing big data for development programs' (Clausen 2016). But Twitter is only one of many platforms which the initiative aims to involve in its vision of data philanthropic, public-private collaborations for development (Kirkpatrick 2016). Humanitarian initiatives such as the Ushaidi Haiti Project (UHP) also gained significant insights into which and where medical support and aid was needed in the aftermath of the 2010 Haiti earthquake. It did so by analysing a variety of (non-)digital sources. UHP established a digital map, bringing together: geographically located tweets; SMS sent to an emergency number; emails; radio and television news; phone conversations; Facebook posts and messages; email list-contributions; live streams and individual observation reports (Meier 2015, 2ff.).

Privacy concerns regarding data retrieved from Twitter, as indicated above, are commonly seen as unreasonable. Still, there are researchers who have stressed users' expectation of privacy even under these conditions (Zimmer and Proferes 2014). But how do we know how users perceive and are affected by research using their data, given that informed consent is neglected and other qualitative data on the issue are still largely missing? This issue becomes even

more complicated when looking at social networking sites and content for which the differentiation between public and private is more ambiguous, as in the case of Facebook 'likes' and other digital interaction data.

Unhealthy Likes: Data Retrieval Through Advertising Relations

Social media data are not always as accessible as in the case of Twitter. In some cases, big data access is granted exclusively or under more restrictive conditions. Researchers who intend to use such data need to acquire access in ways defined by the respective platforms and the corporations that own them. This has been achieved by establishing private-public partnerships, that is: collaboration between employees (potentially researchers) of tech corporations and academics working at universities or public health institutions.

For example, platforms such as Google Flu Trends have been based on collaboration between scientists from the United States CDC and Google employees (Ginsberg et al. 2009)[67]. Similar research using Yahoo search queries as data for influenza surveillance involved a Yahoo Research employee (Polgreen et al. 2008). The first mentioned author of the 'emotional contagion experiment' (Kramer, Guillory, and Hancock 2014; see *Informed Consent* in chapter 3 of this book) works for Facebook's Core Data Science Team. It has been discussed already that the conditions for establishing such partnerships are largely opaque. They depend on corporate preferences and individual negotiations, often in favour of well-known and networked elite universities.

As an alternative to such collaboration and institutional dependencies, researchers have explored a form of data access which allows for possibilities comparable to the described Twitter data: they place themselves in the position of advertising customers. This does not necessarily mean that they pay for retrieved data, even though this has also been the case. Either way, researchers do collect such data via channels originally designated for advertising and marketing purposes. One of the earliest examples of this is an approach which Eysenbach called the 'Google ad sentinel method'. The epidemiologist was able to demonstrate '[…] an excellent correlation between the number of clicks on a keyword-triggered link in Google with epidemiological data from the flu season 2004/2005 in Canada' (Eysenbach 2006, 244). But obviously such data were and are not openly accessible.[68]

Eysenbach described his approach as a 'trick' (245), since the actual Google search queries were not available to him. Instead, he created a 'Google Adsense' commercial campaign, which allowed him to obtain insights into potentially health indicative data. His method was not able to obtain actual search query quantifications, but he was able to factor in those users who subsequently clicked on a presented link. When (Canadian) Google users entered 'flu' or 'flu symptoms', they were presented with an ad 'Do you have the flu?', placed by

Eysenbach. The link led to a health information website regarding influenza. As an alleged advertising customer, Google provided the researcher with quantitative information and geographic data for users who clicked on the placed ad. When relating these clicks to data from the governmental 'FluWatch Reports' (provided by the Public Health Agency Canada), he detected a positive correlation between the increase of certain search queries and influenza activities. Eysenbach describes his approach as a reaction to a 'methodological problem [which] lies in the difficulties to obtain unbiased search data' (2006, 245). The ethical implications of this method and of the conditions leading up to its development are up for debate, however.

The use of data meant for advertising customers has been comparatively less common, and was predominantly applied to North American users. Research involving Facebook's social data is noteworthy. Advertising on Facebook has been used for recruiting study participants (Kapp, Peters, and Oliver 2013). In such cases, researchers had to pay for the placed ads and received, in addition to responses from interested individuals, access to the data generated in this process.[69] However, scientists have also registered as business customers for Facebook's advertising and marketing services – which disclose some data freely, without any necessary payment. Based on the latter approach, Chunara et al. (2013) and Gittelman et al. (2015) explored how Facebook's developer platform, available APIs and data may be utilised as means of public health surveillance.

In terms of relevant actors, it makes sense to first look at the specific stakeholders involved in both papers. The paper by Chunara et al. (2013) is based on collaboration between academics working at US universities. The team consulted an (unspecified) advertising company for information on Facebook's data retrieval possibilities and conditions.[70] Gittelman and his co-author Lange were/are (in 2017) both employed at Mktg, Inc. which presents itself as 'lifestyle marketing agency'. Gittelman is the company's 'president CEO'.[71] Further co-authors are employed at the CDC (National Center for Chronic Disease and Health Promotion) and USDA National Agricultural Statistics Service. These constellations are an insightful indication of the expertise needed and merged in such research.

Expertise in big data analytics has been extensively cultivated in marketing and advertising contexts. Related actors possess skills which are crucial for employing social media data. This has enabled them to participate in research involving big data, complementing the expertise of researchers specialised in, for example, public health. In these contexts – involving public-private collaboration or consultancy relations – marketing expertise becomes an asset in public health research. On the side of the users, it also means that Facebook content posted, exchanged or clicked on for entertainment purposes and social interaction is turned into health relevant information. In this case, Facebook users whose data were retrieved for relevant studies are particularly crucial

stakeholders. In both abovementioned cases, as noted earlier, these are users located in the US.

Chunara et al. (2013) assess how various Facebook data may contribute to public health surveillance of obesity. According to the authors, the availability of geographically specific data makes the social network a particularly valuable source. Facebook allows potential advertising customers to pre-assess and choose potential target groups '[…] based on traits such as age, gender, relationship status, education, workplace, job titles and more.'[72] This specifically includes information on geographical location, interests (e.g. hobbies or favourite entertainment) and behaviours (e.g. purchase behaviours or device usage). Through Facebook for Developers and its advertisement/marketing platform, such data were accessed by Chunara et al. (2013). As the authors describe:

> 'The platform provides the number (found to be updated approximately weekly) of users who fall under the selected categories and demographics at the resolution of zip code, city, state, or country including surroundings at varying geographic radii. Categories are determined through individuals' wall postings, likes and interests that they share with their Facebook friends and through which they create a social milieu.' (Chunara et al. 2012, 2)

Categories can be accessed as aggregated user profiles, based on certain areas of indicated interests and habits. Chunara et al. selected particularly the categories 'health and wellness' and 'outdoor fitness activities' as relevant indicators to assess obesity prevalence. Social media data focused on these categories was then related to data from the CDC's Behavioral Risk Factor Surveillance System. The authors found '[…] that activity-related online interests in the USA could be predictive of population obesity and/or overweight prevalence' (Chunara et al. 2013, 6). While the authors do not present this as a surprising outcome as such, they frame their study as a contribution to identifying viable, novel methods and complements in public health surveillance. Potential limitations are discussed carefully (ibid, 4-6); however, these are depicted as methodological challenges rather than reasons for ethical concerns.

The abovementioned study involves diverse social data sources, for example content such as wall postings, likes and indicated interests. In comparison, Gittelman et al. (2015) focus on 'likes', i.e. users' clicks on Facebook's famous like-button. This button is predominantly read as an expression of interest in as well as support and sympathy for certain content. The authors examine how the data emerging from users' 'liking' of content may act as potential health indicators for mortality and disease rates, as well as so-called lifestyle behaviour. Comparable with the approach of Chunara et al. (2015), they use aggregated data of users, sorted by zip code. These users 'liked' certain items, falling under certain categories.

The data are retrieved through Facebook's marketing/advertisement platform for developers. As the authors explain, they selected three main categories from the available eight overarching categories – events, family status, job status, activities, mobile device owners, interests, Hispanic[73], and retail and shopping – relevant to US audiences. They chose 'activities' and 'interests' because these include the sub-categories 'outdoor fitness and activities' and 'health and well-being', as assumed factors for self-care and physical activity. The category 'retail and shopping' was selected as an indicator of socio-economic status (SES), which is linked to health- conscious behaviour and financial opportunities to realise a healthy lifestyle (Gittelman et al. 2015, 3).

The data obtained from Facebook were then correlated with public health data from the US National Vital Statistics System (e.g. on mortality rates, disease prevalence, and lifestyle factors) and the US Census, as well as self-reported data from the Behavioral Risk Factor Surveillance System (BRFSS). The latter includes information on habits such as smoking and exercise, the health insurance status ('insured'), and health conditions such as diabetes, prior heart attack or stroke. Based on correlations between these sources and social media data, the authors argue that, in combination, Facebook likes and socio-economic status (SES) indicators, for example income, employment, education information, can predict the tested disease outcomes (see Gittelman 2015, 4). Moreover, they stress the behavioural significance of such data by portraying 'likes' 'as a measure of behaviour' and determining 'the behaviors that drive health outcomes.' (ibid).

In this sense, Facebook data are not merely presented as indicators of existing health conditions, but also of likely, future behaviour. The latter assumption, in terms of technological promises, reduces the complexity of health-relevant behaviour to schematic categories which have been conceptualised for adver-tising purposes.[74] Moreover, it does not take into account the fluidity of social media as such – which has been demonstrated, for instance, by Facebook's 2016 introduction of 'like' alternatives called 'reactions'.[75] As opposed to emphasised, ambitious promises, an ethics section and reflections on eventual moral con-cerns are entirely missing from Gittelman et al.'s (2015) article.

For the US, access to health relevant information via social networking sites such as Facebook is possible due to the lack of legal frameworks protecting users' rights to certain big health data. With regards to medical privacy, the Electronic Frontier Foundation (EFF) stresses that social networking sites and other online services pose severe risks and threats to individuals' control of personal data. This applies particularly to users located in the US. The EFF details this situation and its implications as follows:

> The United States has no universal information privacy law that's com-parable, for instance, to the EU Data Protection Directive. [...] The baseline law for health information is the Health Insurance Portability

and Accountability Act (HIPAA). HIPAA offers some rights to patients, but it is severely limited because it only applies to an entity if it is what the law considers to be either a 'covered entity' – namely: a health care provider, health plan, or health care clearinghouse – or a relevant business associate (BA). This means HIPAA doesn't apply to many entities who may receive medical information, such as an app on your cell phone or a genetic testing service like 23andMe (Electronic Frontier Foundation n.d.).

This also implies that US users' Facebook or Twitter data, despite their actual use as health indicators, are so far not protected under HIPAA.[76] In Europe, the data protection directive mentioned in the above quote has been meanwhile replaced by the EU General Data Protection Regulation (GDPR). A directive sets out objectives to be achieved by all EU countries; in contrast, a regulation is a legally binding legislative act. The GDPR was adopted in April 2016 and will be fully implemented by the end of May 2018 (see also Morrissey 2017). Its consistent application across the EU will be overseen by the European Data Protection Board (EDPB). The GDPR has been described as an important step towards safeguarding European users' rights and privacy in a global data economy. At the same time, businesses have been concerned about compliance requirements and practical challenges, implying economic disadvantages. Moreover, in response to earlier/draft versions of the GDPR, biomedical researchers, notably epidemiologists, raised the issue that parts of the regulation allow for interpretational leeway and could lead to overly restrictive informed consent requirements (Nyrén, Stenbeck and Grönberg 2014, 228ff.).

As so often, data protection turns out to be negotiated as a trade-off between public wellbeing and broader benefits, a society's capacity for innovation, and individual rights. With regards to Europe, tensions between users' rights and data as a driver for innovation have been extensively considered in documents released by the European Data Protection Supervisor (EDPS), an independent EU institution. It has been pointed out, from an innovation and research perspective, that the legal restrictions implemented in this field may impede the productivity of research and innovation.[77] Even in the EC General Data Protection Regulation (GDPR) exceptional status is granted to the use of personal data in certain situations, referring to the need of weighing the public good and individual rights:

> 'Such a derogation may be made for health purposes, including public health and the management of health-care services, especially in order to ensure the quality and cost-effectiveness of the procedures used for settling claims for benefits and services in the health insurance system, or for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.' (The European Parliament and the Council of the European Union 2016, 10; §52)

Such retrenchments are necessary; they open up possibilities for highly relevant and needed health research. But they also require further ethical considerations on the question of which cases derogations are reasonable. It is therefore problematic if ethical reflections on these issues are neglected when it comes to big data-driven health surveillance. Such research fails to address, in terms of normative rightness, why informed consent appears dispensable under certain conditions. This also means that public debate on this issue is a priori uninformed, a factor that is indispensable for the formation of valid social norms, according to Habermas.

One of the normative arguments recurring in various big data-driven health studies and justifications is the emphasis on the 'cost-effectiveness of the procedures' (see also the quote above). Already in his early 2006 study, Eysenbach stressed the timeliness and accuracy of what he called the 'Google Ad Sentinel method'. He also pointed out its cost effectiveness compared to more traditional approaches to influenza surveillance (Eysenbach 2006, 244; see also 246). Similar statements can be found in the papers by Gittelman et al. (2015) and Chunara et al. (2013). Gittelman et al. describe their method as a contribution that 'directly affects government spending and public policy' and comes at 'a fraction of the cost of traditional research' (2015, 7). Chunara et al. (2013) stress that their big data-driven research offers 'a real-time, ease of access, low-cost population-based approach to public health surveillance' (2013, 6).[78] This emphasis on financial benefits needs to be seen in the context of health care systems which are under ever increasing pressure to economise and reduce costs (Kaplan and Porter 2011). The authors strengthen the (misleading) assumption that big data provide a solution for this issue. This conclusion needs to be urgently mitigated by re-emphasising the societal costs looming due to an inordinate, naïve reliance on technological promises, promoted by internet and tech corporations. These costs are related to public health monitoring platforms modelled on fluid big data economies (see the previous sub-chapter on Twitter data); a conceptualisation of users' as static, non-reflective entities; and a negligence of algorithmic biases and recalibration needs.[79]

Public Health and Data Mashups

The studies mentioned and described above, involving Facebook, Twitter, and Google data, have in common that they initially focus on stages of methodological exploration. The authors examine how available data could be analysed and used for public health surveillance. Ultimately though, in most cases, such investigations strive for technological utilisations of their methodological insights. Most of them have a concrete development aspect.

This is obvious in the case of projects funded as part of the US BD2K grant scheme 'Targeted Software Development', which applies to Wang et al. (n.d.) and Young et al. (n.d.). As part of the interconnected projects, the two PIs are

involved in creating a platform which automatically retrieves, analyses, and visualises Twitter data indicative of HIV/AIDS high-risk behaviour (Wang et al. n.d.). Also, the UN Global Pulse Labs are developing practical applications such as the 'Haze Gazer', a crisis analysis tool supported by the government of Indonesia,[80] and implementing public dashboards for, among other uses, 'Monitoring in real time the implementation of HIV mother-to-child prevention programme' in Uganda.[81] Chunara and Brownstein, both of whom contributed to the aforementioned paper on monitoring obesity prevalence through Facebook data (Chunara et al. 2013), are part of a team engaged in various explorations and practical applications of 'digital disease detection' (Brownstein, Freifeld and Madoff 2009). In interdisciplinary collaboration with biomedical and computer scientists, they notably developed a platform called HealthMap. This has been described by *Wired* as a manifestation of Larry Brilliant's wish and vision for a freely accessible, online, and real-time public health surveillance service (Madrigal 2008; see also Chapter 4).

HealthMap is an example of data mashups, which are increasingly common. These are websites which select and combine data from diverse online sources (Crampton 2010, 25ff.). In the case of public health surveillance services, they are often combined with geographic maps. Cartographic visualisations facilitate epidemiological insights into the spatial patterns and spreading of infectious diseases. Maps may support public health professionals in assessing how quickly a disease spreads and which spatial patterns emerge. At the same time, they serve as accessible tools for communicating disease information to the public. Spatial analyses and visualisations of epidemics are part and parcel of public health surveillance (see also Ostfeld et al. 2005).

Already in the mid-1990s, Clarke et al. examined the potential use of emerging Geographic Information Systems (GIS) (i.e. locative processing and visualisation tools) in epidemiology. The authors stressed the promises coming along with such developments: 'GIS applications show the power and potential of such systems for addressing important health issues at the international, national, and local levels. Much of that power stems from the systems' spatial analysis capabilities, which allow users to examine and display health data in new and highly effective ways.' (Clarke et al. 1996, 85) The use of data map mashups is a continuation of previous public health surveillance practices, but opens up novel possibilities and challenges.

The use of data map mashups for public health surveillance has been explored since the mid-2000s. The public services EpiSPIDER (Tolentino et al. 2007; Keller et al. 2009) and BioCaster (Collier et al. 2008) mapped data retrieved from various online sources, such as the European Media Monitor Alerts, Twitter, reports from the US CDC and the WHO. The selected information was then presented in Google Maps mashups. Google Flu Trends (GFT, Ginsberg et al. 2009) can be considered Google's in-house solution for Brilliant's vision of an online disease surveillance system. (Brilliant was involved in the project and paper himself). While GFT aimed at predicting influenza intensities based

on search queries that were previously correlated with traditional health surveillance data, HealthMap's objectives are more diversified in terms both of the diseases included and the data. In both cases though, the retrieved and selected data are/were[82] presented in an interface integrating Google Maps.[83]

For the creation of HealthMap, epidemiological expertise, data science, and bioinformatics had to go hand in hand. In terms of directly involved stakeholders, the platform was developed by interdisciplinary teams of epidemiologists, computer scientists (particularly bioinformaticians), and data scientists. It was launched in 2006, enabled by research from an interdisciplinary team at Boston Children's Hospital, with epidemiologist Brownstein and computer scientists and biomedical engineer Freifeld in leading roles. The project has been extensively documented by involved scientists in publications in leading academic journals (see e.g. Brownstein et al. 2008; Freifeld et al. 2008; Brownstein and Freifeld 2007).

HealthMap received funding from multiple corporations, for example a grant of $450,000 by Google's 'Predict and Prevent' initiative as well as from Unilever, Amazon, and Twitter, and foundations such as the Bill and Melinda Gates Foundation and the Skoll Global Threads Fund. It was also provided with financial support from governmental agencies such as the US Defense Threat Reduction Agency (DTRA), the CDC, the NIH National Library of Medicine, and the Canadian Institutes of Health Research (CIHR). Visually, the interface is dominated by Google Maps: in this map, health relevant information – such as news items on disease outbreaks or tweets concerning disease developments in a certain region – are located. The selection process is automatised, in that certain sources are monitored by default and it is algorithmically determined which content will be included. Depending on the website users' location, a smaller text-box on the right indicates potential 'Outbreaks in current location' which are clustered into twelve disease categories. [84]

HealthMap combines data which are retrieved by scanning multiple sources. Among them are the commercial news feed aggregators Google News, Moreover (by VeriSign), Baidu News and SOSO Info (the last two are Chinese language news services), but also institutional reports from the World Health Organisation and the World Organisation for Animal Health, as well as Twitter.[85] The platform utilises global sources and is not limited to a particular country. These are authored by public health institutions or news outlets/journalists. Before being published, such sources are commonly subject to selection and verification processes during which their quality and correctness is assessed. This applies particularly to organisations such as the WHO, but is also the case for quality journalism outlets (Shapiro et al. 2013). In contrast, microblogging platforms such as Twitter also contain information from individual users. Although this latter source of information may be more current, it is also more difficult to verify (Hermida, 2012). Apart from automatically retrieved social media content, users can also send individual reports: this can either be done through the website's 'Add alerts' function (which is part of the

top menu), by email, text message, phone call (hotline), or by using the mobile app *Outbreaks Near Me*.

News items are a particularly dominant type of data, mostly retrieved from news aggregators, with Google News items being especially prevalent. Therefore, being included in such aggregators enhances the chance for (health-indicative) news items to be presented in HealthMap. These aggregators, maintained by global tech corporations, play an important role as gatekeepers, defining in- and exclusion. In this sense, research concerning the gatekeeping function of such aggregators is highly relevant to projects such as HealthMap, and may be used to assess the implications of such an approach (Weaver and Bimber et al., 2008).

While drawing on news aggregators seems to be a technically feasible/preferable solution, this approach raises questions regarding the selection criteria relevant to utilised big data sources. The presented data go through multiple forms of automated selection: first, they are defined by, for example, the Google algorithm that determines more generally which sources are included in its *News* service. Second, they are subject to an automated process in which the HealthMap algorithm selects information which is considered relevant for disease detection.

In combination with the funding the project received, the used content poses questions regarding eventual conflicts of interests and emerging dependencies. Exaggerating somewhat, technology editor Reilly (2008) remarked of HealthMap: 'We can't officially call the program Google Disease(tm). But that's essentially what HealthMap is.'[86] In an interview, Google 'Predict and Prevent' director Mark Smolinski commented on HealthMap and the decision to provide funding: 'We really like their approach in that they are trying … a really open platform,' […] 'Anybody can go in and see what kind of health threats are showing up around the world' (Madrigal 2008).

The fact that Google material is being used provides the corporation with positive public exposure. It links the company's (branded) content to technoscientific innovation as well as the well-established perception that public health surveillance is an important contribution to societal wellbeing. Whether the use of Google data is, methodologically speaking, the ideal approach for HealthMap remains to be explored. Ethically, the emerging dependencies may result in stakeholder constellations between data providers and scientists which affect future decision making. This latter effect has already been described with regards to pre-big data industry funding (Lundh et al. 2017; Bekelman, Li and Gross 2003).

The dominance of Google News items in large parts of Europe and the US is also likely related to a main methodological challenge already addressed by the scientists involved in the creation of HealthMap. With regards to the used 'web-accessible information sources such as discussion forums, mailing lists, government Web sites, and news outlets', Brownstein et al. (2008) state that '[w]hile these sources are potentially useful, information overload and difficulties in distinguishing 'signal from noise' pose substantial barriers to fully utilizing

this information'. This concern refers to the challenge of selecting relevant data, but it should also be seen in the context of different data sources providing varying amounts of data.

Considering Google News' extensive, ongoing data collection and capacities, sources which provide quantitatively less input run the risk of being overlooked – in this case not by the algorithm, but by those users trying to make sense of visualised data. What is happening here can be (structurally) compared with the common experience of a Twitter user who starts following very vocal corporate, political or governmental account, for example. The constant 'noise' of such quantitatively dominating actors is likely to impede one's perception of other, relevant information sources.

Dependencies and potential conflicts of interest concern the content which is mapped, but also the Google map itself. The fact that content is placed in Google Maps also raises issues concerning sustainability, similar to those dynamics described for Twitter data. Critical geographers were also among the first to tackle the sensitivity of big data and locative information (Dalton and Thatcher 2014; see also Chapter 1). They have cautioned against uncertainties when relying on corporate services in *neogeography*. The latter notion implies that maps are created and processed by actors who are not trained cartographers, but participate in map-making with the help of cartographic online services (see also Rana and Joliveau 2009, 79).

There are various mapping services, such as Google (My) Maps, the abovementioned Ushaidi platform, or the free and open source project Open Street Map, which enable non-cartographers to map information or even to create cartographic surfaces. It has been highlighted, though, that these participatory mapping approaches are still subject to regulations defined by the map hosts. This is especially relevant in cases where the cartographic material is owned by corporations such as Google. Various authors have challenged optimistic assumptions of a 'participatory mapping culture' and its democratisation. They point out that neogeographic practices are defined by access to the internet and digital content as well as digital skills and literacy.

Haklay (2013) criticises the starry-eyed promise that neogeography 'is for anyone, anywhere, and anytime'; instead, the author argues that looking at the actual practices exposes sharp divides between a technological elite and 'labouring participants' (Haklay 2013, 55).[87] In addition to such issues of accessibility and expertise, there are new forms of dependency which are related to the dominance of global media corporations: 'One of the more curious aspects of Neogeography is the high dependency of much activity on the unknown business plans of certain commercial bodies providing API's for mapping.' (Rana and Joliveau 2009, 80) This also has an influence on the sustainability of projects relying on commercial APIs, since the conditions for using them may change – as also remarked with regards to prior research approaches.

Potential conflicts of interests and dependencies in big data-driven health projects should be placed in the context of broader ethical considerations for

datafied societies. Calling attention to seminal changes emerging in research connected to global tech corporations, Sharon (2016) argues that since unfolding '[…] power asymmetries may affect the shaping of future research agendas, they deserve greater critical attention from medical researchers, ethicists and policy makers than is currently the case.' (564).[88] It is striking that such concerns are rarely an integral part of techno-scientific explorations of big data-driven public health research. What can be considered 'disruptive communicative action' in Habermasian terms does occur, for instance in those critical contributions which I have continuously referenced above. But these disruptions are never moved toward a level of 'higher' argumentative discourse.

An engagement with ethical issues that takes the side of those involved in big data-driven public health surveillance is reduced to justifications of research practices, or in some cases is even missing. In those, still exceptional cases, where such validity claims to normative rightness are raised and challenged, a discursive divide between those arguing from an ethical and those from an innovation-driven, methodological perspective prevails. Ethical arguments appear to unfold in distinct spheres rather than in actual dialogue. From a discourse ethics perspective, this is problematic, since it weakens the validity of social norms and moralities crucial to respective research approaches. In the following, final Chapter 6, I will elaborate on this conclusion by tying it back to the critical perspectives and theory introduced in Chapter 2.