

Data as environment, environment as data: One Health in collaborative data-intensive science

Big Data & Society
 April–June: 1–13
 © The Author(s) 2024
 Article reuse guidelines:
sagepub.com/journals-permissions
 DOI: 10.1177/20539517241234275
journals.sagepub.com/home/bds



Lucilla Barchetta  and Roberta Raffaetà 

Abstract

This article analyses the operationalization of One Health in the context of data-intensive science in response to the COVID-19 outbreak. Building on ethnographic field research and revisiting the lives of a knowledge infrastructure of interdisciplinary collaboration set up online in the early phase of the COVID-19 health emergency, the article develops the notion of “data as environment.” This environment is a contact structure that entangles knowledge systems, subjects, processing tools, and mediated bio-socialities in processes of data-intensive knowledge co-production. Claims for new collaborative approaches between the biomedical, environmental, and social sciences are increasingly marked by the emergence of digital knowledge-making infrastructure that leverages data, knowledge, and expertise from different disciplines and sectors to increase scientific productivity via data-sharing technologies. Yet, digital knowledge-making infrastructures appear self-evident when they are in place, while data are often conceived as inert and disembodied information units separated from social relations of research. The argument that data are an environment expands anthropological thinking on data and digital knowledge-making infrastructures by enlightening political-ethical questions that are at stake in the emerging technoscientific worlds of the Anthropocene.

Keywords

One Health, data-intensive science, ethnography, knowledge-making infrastructures, data, environment

Getting hands on data, getting hands on the world

This article¹ presents an ethnographic study of One Health (OH) and its operationalization in a digital infrastructure of interdisciplinary collaboration. OH has been defined as an integrated, unifying approach that aims to sustainably balance and optimize the health of people, animals, and environments as closely linked and interdependent systems (One Health High-Level Expert Panel et al., 2022). Since the early nineteenth century, the OH approach developed as a result of social, epistemological, institutional and political factors that fostered integrated, collaborative approaches to health across the Atlantic (Woods et al., 2018). Yet, the nature and extent of the current ecological crisis and the recent COVID-19 pandemic demonstrate the urgency to address the reciprocal relationship between health and environment, by operationalizing OH into policies and concrete actions. The OH vision calls for an understanding of health that comprises the planet’s wholeness. The approach is, therefore, inherently interdisciplinary as it aims at mobilizing knowledge, data and

expertise from multiple sectors and disciplines at different scales to tackle health threats and foster well-being.

A growing body of literature recognizes the importance of data technologies for operationalizing this conceptual move by sharing and integrating data across different sources and analyses (Benis et al., 2021; Farman and Rottenburg, 2019). The OH vision depicts a world in which the emergence of new alliances among different actors, the use of cloud computing technologies and open science (OS) tools—such as visualization, open databases, journals and platforms for data sharing—and the implementation of data management protocols will radically advance scientific knowledge production about health across

Department of Philosophy and Cultural Heritage/NICHE, Cà Foscari University of Venice, Italy

Corresponding author:

Lucilla Barchetta, Department of Philosophy and Cultural Heritage/NICHE, Cà Foscari University of Venice, Malcanton Marcorà, Dorsoduro 3484/D, 30123 Venice, Venezia, Italy.
 Email: barchlucil@gmail.com

disciplines and scales. Yet, in approaching OH as a socio-technical apparatus, data technologies should be considered not only as a technical tool to achieve the OH vision but also as the causes and drivers of OH popularity and institutional support in this historical moment. Our article shows this tension between the search for a systemic response to ecological predicaments and technological aspirations. Moreover, it also documents that the transition from OH vision to technological application may be not as smooth as assumed. Okune and colleagues (2018: 2) write, “There is an assumption that once these virtual infrastructures are in place, researchers and other collaborators will be able to participate in the creation of scientific knowledge in more equitable and efficient ways.” As this paper illustrates, this is hardly the case. Some OH projects on digital knowledge-making infrastructures narrowly consider data as commodified units of information that can be shared when a technological apparatus is in place and ready to be used by a scientific community of peers. In contrast, we argue that data, far from being fixed and enclosed entities, are an environment. To think about data as an environment means to think of data as a contingent contact structure that interrelates—via use—technologies, human subjects, knowledge systems, mediated biosocial components, and socio-economic structures.

Conceptualizing data as an environment comes out of our ethnographic work within the field of OH: the environment is a foundational idea in the OH vision, and the scientists with whom we collaborated sought to integrate an environmental perspective into their analyses of health issues. As anthropologists, we take seriously the method of “thinking through things” (Henare et al., 2007) and considering an ethnographic “substance” as a method (Dumit, 2021). In the case of our research into OH, the concept of environment became an ethnographic substance: it appeared in daily conversations and thus it also molded the shape of our thought. It is strongly ethnographic also in the sense that it has been animated by our commitment to intervene in and interfere with our site of observation and, in this light, conceptualizing “data as environment” functioned to facilitate communication with our collaborators for whom the environment was an object of both affection and cognition, and more so than other abstract terms.

Framing “data as environment” also relates to a broader field of scholarly debates, and is inspired in particular by scholars who established a connection between ecological metaphors and technology. For example, Haraway offered early insights into how the advancement of molecular biology and genetic coding translated ecosystems, as immune systems and biotic components, from objects of knowledge into information-processing devices (1987: 19–20). Star and Ruhleder (1994) took the ecological framework of information spaces to recognize the material and political implications of a digital collaborative infrastructural tool designed for a geographically dispersed

community of geneticists. Those works anticipated today’s information technology jargon, in which “IT environment” is frequently used to indicate all the technology components (hardware, software, and networking) that serve the needs of both developers and users of such infrastructures. “Data as environment” also expresses solidarity with more recent scholarship using environmental metaphors to describe information spaces (Delfanti, 2013; Edwards et al., 2013; Gabrys, 2016). We draw and expand upon these approaches to probe the empirical and conceptual utility of thinking of “data as environment” as a structure of contact that comprises and simultaneously exceeds computational processes.

An important source of inspiration for this article have been various authors who have written about “data environments,” or the circumstances under which data are handled, giving rise to concepts such as “data infrastructures,” “data ecosystems,” “data communities,” “data journeys,” “data collaboratives,” and “data cultures” (Borgman, 2012; Bowker, 2000; Leonelli and Tempini, 2020; Poirier and Costelloe-Kuehn, 2019; Ribes, 2006; Verhulst, 2015). Leonelli, in her foundational work on data-centric biology (2016), contributed an important relational conception of data. Building on this, “data as environment” emphasizes the strongly entangled and ontoepistemic character, in Barad’s (2007) words, of the scientific practices examined in our case study. Barad, both physicist and philosopher, observes that assuming “something” is “in relation” with “something else” is to already establish a disconnection between the two. So, to conceive data as influenced by the circumstances under which these are handled, captures important aspects of how data are produced and mobilized but still assumes the existence of something called “data” as distinguished from something called “environment.” Instead, we conceive data themselves as environment, a structure of contact that emerges in putting into relationship various elements, activated by use, maintenance and slippages. In looking at what happens before we have data—how a data come to be data—and after—how data are used—we address the mainstream criticism that data are self-contained bits, enclosed and autonomous objects, extracted and abstracted from the flux of becoming. We show that in their creation data are dependent from social coordination, and traces of this are apparent in their use too. In thinking at “data as environment” we draw on Barad’s concept of “intra-action”:

The neologism “intra-action” *signifies the mutual constitution of entangled agencies*. That is, in contrast to the usual “interaction,” which assumes that there are separate individual agencies that precede their interaction, the notion of intra-action recognizes that distinct agencies do not precede, but rather emerge through, their intra-action. It is important to note that the “distinct” agencies are only distinct in a relational, not an absolute, sense, that is, *agencies*

are only distinct in relation to their mutual entanglement; they don't exist as individual elements. (Barad, 2007: 33, emphasis in the original)

In our fieldwork, indeed, we found that data are not just an inherently existent entity; they extend beyond their material or empirical specificity in a particular moment in time (in the form of a document, Excel file, research manuscript, and number). Data, in the OH practices we analyzed, are more than a network or assemblage of elements and social relations. They are a structure of dynamic, emerging and not linear connections, characterized by ontological and affective excess (de la Cadena and Escobar, 2023).² This excess surpasses the computational realm and is experienced through the perception of conflicting temporalities, in encounters with the unforeseen and the unlooked for, or through tasks that require researchers to double their attempt to keep up with both shifting scientific objects and routine activities. Thinking about “data as an environment” also mirrors the present moment of scientific debate, especially with reference to postgenomics (Meloni, 2014; Niewöhner and Lock, 2018; Richardson and Stevens, 2015). Twenty-first-century genomics does not understand the environment as external to an organism but rather, as Baedke and Buklijas (2023: A6) write, “organisms (co-)construct their environments and, as a consequence, themselves.” In our fieldwork, data were not simply the product of social relations—a common position in social studies of data—but data were those same relations. Indeed, the circumstances in which data are handled are not merely external environments that inform what data are, what they mean, how they shift. Data and the environments they open up—via use—are (co-)constitutive and not clearly distinguishable.

To make sense of the environments opened up by uses of data, digital knowledge infrastructures support scientists in attuning to an increasingly complex and mixed datascape in which the processing of data co-exists with a multifaceted set of open-data policies, workflow-management protocols, computing capacities and approaches to intellectual property. But, we argue, digital knowledge infrastructures should be both stable (providing a framework for action and social coordination) and flexible and dynamic (open to maintenance and use). Therefore, another foundational concept of this paper—slippage—helps us to capture the specificity of data as environment. “Slippage” is a term coming from mechanical engineering that refers figuratively to deviations from the norm: from what we think will happen, what does not unfold exactly as intended, and the frustration of what doesn't happen or is difficult to accomplish. In our case study, slippages manifested as difficulties accessing and understanding data (“slippages of access”), and as the excess of labor linked to the challenges of co-authoring, co-experimenting and co-organizing (“slippages of co-labor”). These slippages,

we found, are an unavoidable infrastructural condition, and they offer an important ethnographic lens to reflect on the relationships between data, science and society. In this article, therefore, we take as a point of departure the combination of multiple perspectives from critical data studies (Beaulieu and Leonelli, 2022; Douglas-Jones et al., 2021; Kitchin, 2014), knowledge infrastructure studies (Edwards, 2017; Edwards et al., 2013; Okune et al., 2018) and OH scholarship (Craddock and Hinchliffe, 2015; Gibbs, 2014; Leboeuf, 2011). Conceptualizing “data as environment” is also part of a wider debate about the anthropological understanding of the very concept of data, in light of increasing demands for FAIR (findable, accessible, interoperable, reusable) data in European funding schemes and research (de Koning et al., 2019; Dilger et al., 2019; Pels et al., 2018). In combining these perspectives, we take “data as environment” as an anchoring point from which to examine efforts to operationalize systemic and interdisciplinary notions of health in a data-centric world. How do scientists handle the tension—a form of slippage—between the assumptions underpinning a multidimensional notion of health and the realities of data sharing for One Health operationalization? How do values, technical restrictions, scientific routines, and beliefs impact how scientists think about and approach data sharing and multidimensional notions of health?

A methodological note on the “We”

This article draws on our experience as two anthropologists in a digital knowledge infrastructures we call Health for All (HfA), set up at the beginning of the COVID-19 outbreak to understand the pandemic's mechanisms and diffusion in an OH framework that sees the virus as not the sole cause of pathogenicity but connected to various elements such as mobility, consumption of drugs, clinical data, and environmental ones. Starting in 2020 and lasting about two years, we conducted semi-structured interviews in person and online with 22 members of the HfA network,³ observed online meetings (as both participants and non-participants), made short-term field visits, and analyzed global health reports and peer-reviewed articles.⁴ Interviews aimed to reconstruct the professional and research biographies of HfA members, the collaboration processes that motivated them to join HfA, their visions and hopes and the main difficulties they encountered in scientific collaboration. In addition, discussions about OH and open science made it possible to analyze scientific subjectivities at a time when scientific and academic research is increasingly competitive and data-driven.

In the field, we visited diverse locations and organizations, notably a private foundation conducting research in the area of data science and digital public health, and an innovation laboratory specialized in plant protection. These organizations are both located in northern Italy and

are two of the most active stakeholders of the HfA network. Field visits provided opportunities to meet individual scientists and further discuss issues previously raised during online interviews. Field visits included a trip to what we call the Innovation Unit of the European Data Centre For Open Science (EDCOS), a very prestigious European science infrastructure that provided computational support and storage space for HfA activities. The HfA members we interviewed were largely connected to the Italian scientific diaspora with backgrounds that ranged from biomedical research and clinical practice to statistics, health economics and plant science. Interviews were conducted in Italian, first transcribed with the software Amberscript, and then manually revised. Field notes, both handwritten and computer processed, and ethnographic doodling complemented these research materials, all of which were then thematically analyzed using an approach inspired by Poirier and Costelloe-Kuehn's (2019) heuristic method for data cultures. The authors' multi-scale heuristic model has seven dimensions spanning from "meta" to "nano," addressing both technological, socio-technical and epistemological aspects of data activities. Rather than employing Poirier and Costelloe-Kuehn's heuristic model as a rigid deductive model, this heuristic gave inspirational input and served as a general guideline for understanding data practices within and beyond our research community in our study. It aided in identifying analytical paths that were linked to the greater corpus of anthropological study on data practices without diminishing contextual distinctness. This approach motivated us to argue for emphasizing concerns about data access and collaboration through data, rather than data scales, as an analytical lens to foster anthropological thinking and generalization that is sensitive to unstable and distributed contexts of collaboration.

With its attention to the intertwined dimensions of the immaterial and the material, ethnographic research has acquired a prominent role in scientific endeavors that examine the social life of cyberspaces, digitalization, and quantification, and their power relations (Horst and Miller, 2013). This research also underscores the role of ethnography in sowing seeds of collaboration in increasingly interdisciplinary research settings (Callard and Fitzgerald, 2015; Pedersen, 2023). Being invited as anthropologists to observe and facilitate collaboration in a professional context made this research possible. However, interactions with specific interlocutors (individual members and teams) were agreed upon on an ad-hoc basis by negotiating access and participation with the network's members. Therefore, access to the field was segmented and intermittent: how it occurred shifted across "fields" and over time also in relation to the unstable lifecycle of HfA. As part of our ethical protocol, which aims for substantive interdisciplinary dialogue, we shared the full text of this article with the member of the network who provided

comments and further insights towards the final version of the text.

Chronicles from a galactic infrastructure

At the beginning of the pandemic, the urge to "make sense" of COVID-19 led to an acceleration in collaborative science, leveraging the strengths and expertise of researchers trained in different fields. This phenomenon is also linked to shifts over the last 40 years in how scientific research is organized, funded and circulated, especially in the biomedical and environmental sciences (Delfanti, 2013: 17; Leonelli, 2023). The emergence of the HfA's digital infrastructure follows this same path, albeit reinforced by the founder's strategic vision and distinctive OS ethos, and enhanced by the computing support of EDCOS. Sara, the founder of HfA, is an Italian veterinary virologist working and living on the West Coast of the United States, though, as she once stated, she always remained a "European scientist." She was deeply involved in the Italian political system before moving to the US. In March 2023, when we were finalizing the first draft of this article, she decided to move back to Italy. Sara's mobility between Italy and US mirrors a larger trajectory, too, with the US as the cradle of OH and Europe as the institutional context to make it flourish. Sara has been a long-term champion of open access and data sharing, and founding HfA represented a further commitment to OS.

During the first COVID-19 lockdown, she illustrated the project to HfA members using the image of a galaxy made of thematic units: planets. At the center of the galaxy stood the sun—EDCOS, the technology hub that indirectly connected all data and actors, however distant or quarantined. This gives an idea of how much hope was assigned to technology in realizing the OH vision. Also, an open environment of collaboration was thought to increase scientific productivity in a time of emergency, when data and its circulation become more and more valuable. As anthropologists, we have been represented as a comet crossing this galactic space of collaboration. Sara imagined HfA as a way to advance the OH vision toward a more holistic, circular vision of health, which she embraced as a foundational concept, on which her long career and publication history rested.

To address the technical needs of network's members, Sara and her team put time and effort into developing a governance framework to organize the legal, financial and socio-technical aspects of a collaborative research environment. However, over the following months, the governance framework was revealed to be unstable. Instead of a wholly coordinated galaxy, the life cycle of HfA has been mostly dependent on how research was organized and delivered within specific thematic units and working groups. In addition, the EDCOS data center lost its prominent place over time. One of the reasons for EDCOS's withdrawal may

have to do with HfA's difficulties in identifying sources of funding and organizational strategies after the first phase of the pandemic, which drove scientists' and society's fears but also hopes. Moreover, EDCOS's waning role was linked to the restructuring of its Innovation Unit, which broke collaborative relationships and shifted EDCOS's strategic focus. But more than chronicle the death of a galaxy or the failure of a project, we wish to depict the efforts and the practical and affective potential that can circulate across a galactic interconnected system of data exchange, and also offer some insights into what is needed in order to achieve it.

Spots of data here and there: Environments of excess mortality

The first country to be hit by the pandemic after China was Italy, and one HfA group sought to understand why Italy, and the Lombardy region in particular, was hit so hard and fast by COVID-19. Consistent with a systemic approach to health, the pandemic emphasized the need to consider the socio-economic context of COVID-19-specific deaths. Our key interlocutors were Matteo (health practitioner) and Sofia (health economist), the operating hands of the group, who worked for the Lombardy Health Research Centre (LHRC), a privately funded research agency founded in 1978 to promote research on healthcare management.

The group began to study excess mortality (EM), which some have dubbed the key metric to understand the COVID-19 pandemic. EM measures the increase in the total number of deaths from all causes—not only deaths attributed to COVID-19—in a given period, as compared to what could generally be expected for a reference period such as, for instance, the average number of deaths in the previous five years. The identification of the Lombardy region as one of the first hotspots of the pandemic in Western Europe gave a characteristic spatial form to COVID-19-related EM and provided a way to engage with the global scholarly discussion on the reliability of EM statistics.⁵ Calculating EM is assumed to facilitate public health regulation schemes in an evolving pandemic because it is a comparable measure. In contrast, average mortality data is less comparable because countries use different criteria to diagnose and record causes of death. Plus, data availability may vary by indicator, country-specific variables (such as significant differences in population) and, more broadly, whether there are reliable diagnostic tests. Making cross-country comparisons thus requires taking several factors into account to interpret pandemic data. Instead, EM is supposed to capture in real-time both confirmed COVID-19 deaths and those that were not correctly diagnosed and recorded by reporting systems. It also aspires to measure deaths from other causes (traffic

accidents, domestic overcrowding, air pollution, mobility, and working patterns) attributable to the pandemic's pressure on poverty, health care delivery, and resulting co-morbidities (diabetes and pre-existing lung conditions). In seeking to provide the net outcome of multiple health and socioeconomic conditions, EM statistics serve to situate COVID-19-related deaths in a socio-economic context, consisting of emergent, porous environments encompassing information infrastructures and intersecting biosocial and urban conditions.

Calculating EM involves aggregating mortality data under the umbrella of “excess deaths” through a ratio or percentage, the so-called *P*-score that measures the deviation from an expected number of deaths in a country. Sofia and Matteo, however, focused on an earlier stage of this calculation: they grappled with the research question of what characterized the data sources used by international agencies to calculate EM estimates during the first wave of the pandemic. Instead of focusing on the statistical calculation of EM, they examined what made it possible for EM to become a reliable statistical measure. The research resulted in a scoping review of EM data sources co-authored by Sofia, Matteo, and other academic collaborators. A table in their paper described sources in columns: name of the source, data provider, data accessibility (public/private), geographical aggregation level (nomenclature of territorial units for statistics, the so-called NUTS), geographical coverage, time coverage, time unit, mortality measurement, sex disaggregated, age disaggregated, and other disaggregation. The table made apparent that death, understood as an individual/biological phenomenon, involves multiple environments—body, medical facility, home, city, region, computer, telecommunication network, and IT system—abstracted and integrated into a database “by which data are packed and circulated” across public health data systems and research institutions (Leonelli, 2016: 5). Critical data studies have shown that when we think about data, we should always be aware that data is never *datum* (Gitelman, 2013). Data are not inherently existent objects. They may become valuable representations of social realities at certain times (what Barad calls “agential cut”⁶) but they are always situated in a flux of becoming, since data are “the result of a journey” across radically different contexts of social coordination, datafication and use (Leonelli and Tempini, 2020). When we deal with data, we should always ask where it came from, how and why it was collected, who created the database and what criteria were selected to classify information. Returning to Sofia and Matteo's inquiry, each national agency database they accessed—the primary, freely accessible source of EM statistics—constituted an abstracted information environment indistinguishable from the agency of researchers. But when EM data are opened up to scrutiny, surveillance technologies, medical institutions, database practices, systems of measurements and labelling, mediated territories and

biosocial components are brought together into interrelationship. The database is where all these components must be rescaled to carry on its journey across other environments: the research's social and IT environment combined with a given type of project.

When computational faith meets data use

HfA research environment is where scientific researchers engage in intensive and time-consuming forms of computational work associated with daily data routines, which entail coordination, cooperation and communication among collaborators. "Computing" means calculating the amount or value of something by using information (esp. numbers) and machines. Since numbers and machines are produced and used by humans, computing systems also coordinate human subjects who interact with data in the socio-material and IT environment developed to conduct a given type of research project. These environments are a constituent part of data. As computational work and the research environment belong together, the results of this encounter cannot be entirely expected because both are embedded in the intersubjective and multimodal sociality of research, which is not limited to scientific components such as methods, research questions and technologies. We interpret this as a further point in our approach to "data as environment" that takes into account interdisciplinary moments of disorientation, pique, bewilderment and confusion that shape the practice of research.

Despite computational faith in EM statistics, recent WHO global estimates suggest that the death toll was considerably greater than the COVID-19-related deaths officially reported (Keusch et al., 2022). EM amplifies, rather than mitigates, the challenge of quantifying the complexity of mortality data, and this outcome is still controversial. EM revealed a paradox that underlies how, in a period of emergency, scientific anxieties about data—how scientific researchers deal with the messiness and uncertainty of digital data—exceed data's capacity to provide certainty (Leonelli, 2021). Anxieties about accessing, scrutinizing and using EM statistics are always part of everyday research environments, that is, the kind of data work researchers handle in their daily scientific routines and interdisciplinary collaborations.

As Sofia pointed out during an interview, building value from the use of national agencies' databases is always "a trade-off" between the legal and technical openness of databases and the granularity of data. "Granularity" refers to how data are detailed: the more open databases are, the more aggregated they are, resulting in less granularity. Granular data have more value than aggregated data for reuse in certain analyses. When opened up and scrutinized, databases may come to life with added value (Ahmed, 2019). Data acquire, in fact, added value through use (searching, inspecting, and scrutinizing), which means

that they become valuable when there is potential to transform data into something else by use, that is, reuse. But to know how to reuse, one has to go through a process of learning through trial and error. Access, thus, is a condition for both use and value, and use also enables "data as environment" to be seen. Data becomes muddled in the research environments of interdisciplinary collaborations, their inherent epistemological uncertainties, and their mundane routines and affective dispositions. In this case of studying EM data sources, it was laborious to screen data sources, both automatically and manually, due to an "explosion" of peer-reviewed, pre-print articles and official reports from Euromomo, Eurosurveillance, the Centers for Disease Control and Prevention, and the World Health Organization—all essential freely accessible mortality-monitoring projects. As Sofia explained, there was an underestimation of the time and resources required to pursue the inquiry:

It turned out to be much more complex than expected, partly because there needed to be a person dedicated 100% to the project. It was one more thing that we were doing out of interest. Moreover, this was particularly crucial because reviewing the literature at a time in history when scientific production exploded made it very difficult to keep up the pace.

Researchers often confront knowledge challenges that cannot be fully anticipated, and for this reason it takes creativity and improvisation to accommodate different capacities and resources to pursue an inquiry, especially when collaborative scientific efforts occur during a devastating pandemic and researchers with different epistemologies are not co-located (Leonelli, 2021). Scientific routines are not always repeated the same way every time, especially when new forms of data work require experimentation (Pink et al., 2017: 3). As Sofia put it, "It was a bit of progressing and finding a way." Her words highlight a slippage, a deviation from what she considered a familiar routine; the project required more time and effort than expected because it involved new procedures, ideas, or activities.

It is precisely such slippages that reveal data as environments, since they shift scientists' attention from the specificity of data features to the socio-political context of data, in this case associated with public health research, the COVID-19 pandemic, and the complex affective states that contoured this research moment. What seemed to her a technical, almost modular task—evaluating data sources—broadened the contextual features of data and provoked ethical-political interrogations about knowledge co-production in the context of pandemic technoscience at a time when researchers felt compelled to produce novel research quickly. Therefore, Sofia felt a sense of relief

when she began to receive support from Francesca, an EDCOS data scientist.

The collaboration between LHRC and EDCOS began when the systematic review was in an advanced stage. With an academic background in statistics and computer science, Francesca started working for EDCOS one year and a half before the start of the HfA initiative. Together with her team of developers, she planned to implement a platform promoted by the Innovation Unit of EDCOS's IT department, in order to expand open-data technologies (software, tools, storage space, and reproducibility service). HfA, and its sub-thematic units, became a "use case" to enhance and test benchmarks in open data management. In particular, Francesca proposed to prototype an algorithmic-based visualization tool for EM in Europe that would have been constantly updated with the newest data streams from processed data sources. The history of the preparatory stage of data processing is stored in a folder within EDCOS's open data repository, containing the GEOJSON file for specific European countries (EU, France, the UK, and Italy)—the open format for organizing geospatial data—and the data normalization records, which in this case refers to the homogenization of dataset features to establish a common geographical scale. However, the data modeling progressed differently from what collaborators planned and therefore was stopped at this earlier phase. It was nonetheless a challenging, time-consuming, "back and forth" process between LHRC and EDCOS. The group's members transferred data sources to Francesca; she then cleaned and homogenized datasets to let EDCOS machines "crunch" data and "feed the visualization tool, once the prototype had been developed." "It was research, and I am not a researcher; I would define myself as a practitioner," claimed Matteo, who joined the project to replace Sofia during her maternity leave. His words expressed feelings of uncomfortableness and uncertainty. He was working as a health practitioner in what he perceived as an academic project and, more broadly, a research practice where he needed to be grounded. In this HfA project, the work made him feel unsettled and caused him to question his ability to make sense of data:

I do less research, I am not doing a PhD. [...] We do mostly consulting rather than research; then, there is a research part in it, but my background is not even rooted in methods, much more on practitioners. I am very much at the first experience, and my contribution is minimal.

After clarifying where the difficulty was found, he described data access as a "not extremely easy" process, because every national data source has its own statistical conventions and definitions. "There are a number of challenges that one finds just by doing them or being an expert in building datasets," he added. His words highlighted that HfA assembled different computing capacities,

often clustered in specific methodologies that reflected the institutional computing history of the research organizations involved in the collaborative effort. The diversity and uneven distribution of computing experience across groups and individual members became evident—another form of slippage—but such diversity and unevenness were left open and unresolved, slowing down and dysregulating the collaborative process. The technical burden of homogenizing data was left mainly in Francesca's hands. The lack of a coordinated and stable framework regarding how to distribute labor, capacities and resources undermined the researchers' capacity to proactively engage with their differences in training, to co-learn tasks from each other, and ultimately to understand the complexity of the social, affective and epistemological grounds on which researchers encountered one another. In the end, co-labor resulted in an excessive discharge of labor. Francesca found herself simultaneously performing the roles of the data expert, with the cultural expectations that this status elicits in terms of efficient knowledge; the data scientist, whose knowledge about data science is perfectible; and the technology mediator, expected to bridge technologies with their local interpretation and application.

Slippages and interdisciplinary knowledge production

The life of HfA demonstrated that slippages in digital knowledge infrastructures are an essential infrastructural dimension for seeing the emergent and entangled environments that the data uses radiate. The degree to which slippages become a focus of attention, however, is dependent upon the capacity to develop an informed and inclusive approach to understanding data and attending to the complexity of the social relations in which researchers operate. To recognize "data as environment" is always the object of an intersubjective interchange that asks scientific researchers to look at and beyond multiple screens. As a result, "slippage" refers to more than the technicalities of access and use; it also concerns attending to the links and efforts to coordinate multiple environments. Hence, slippages are co-constitutive of digital knowledge infrastructures because they allow for seeing data through an environmental perspective encompassing the socio-technical, interpersonal, affective, infrastructural, political-ethical and onto-epistemological aspects of doing collaborative science. Co-labor in interdisciplinary spaces entails a choreography of activities—passing objects from screen to screen, from software to software, from keyboard to keyboard—which unveils how "data as environment" unfolds via this choreography of uses and comes to be rescaled from a Zoom room into a database, and from a database into a machine-readable dataset.

Reusing EM data sources published by national and international agencies requires converting them into a machine-readable format that may vary by data type. One

of these formats is CSV, whose acronym means “Comma-Separated Values,” a standard format for spreadsheet data. In a CSV file, data is represented in a plain text file, with each data row on a new line and commas separating the numerical values on each row. Francesca planned to systematize data sources into multiple CSVs by separating information into a tripartite spreadsheet organization of, respectively, qualitative description (first spreadsheet), quantitative information (second spreadsheet), and results (third spreadsheet). The making of CSV files constituted an additional preparatory stage of data processing that she was unable to complete. The collaborative focus on this preparatory stage was the spark that brought us to think, in this ethnographic scene, of data as an environment. To paraphrase biologist and writer Kriti Sharma (2015, pp. 36–37), the making and validation of knowledge—in our case, knowledge regarding EM and an open-data format—are intricate processes none of which point to data as inherently existent objects, but all of which point to how data come into being entangled in processes of social coordination.

Understanding what using open-data technologies entails outside data-science circles, like research in general, is rooted in learning what inhabiting an interdisciplinary collaboration is like and why it is often hard to accomplish. The opportunity to meet Francesca in person helped us better grasp the context of the use of open-data technologies. In May 2022, Francesca moved into her new office on the second floor of the EDCOS data center. The topography of EDCOS is vast: people usually move by car or bike, and walking is rare, except for short trajectories (office to restaurant and office to another office). Walking the route from Reception to Francesca’s office was like passing through a maze of light blueish and grayish flat buildings characterized by a predominant 1980s architectural style. As we walked in the sunlight, the background noise of power and cooling machines revealed our spatial proximity to the data center. Francesca said: “It’s difficult for people to understand the role of the Innovation Unit. Results are usually intangible and become visible on a long-term basis. There is no way to measure it before it actually produces an impact.” She then mentioned Kuhn’s seminal work, *The Structure of Scientific Revolutions*, to emphasize the slow pace of technological innovations, which sometimes doesn’t allow for capturing the mundane shifts that produce scientific advancements in research environments. She detailed how EDCOS planned to upgrade the ICT infrastructure in light of emerging needs in data acquisition, storage, computing platforms, networks and communication, and data analytics. This process, she explained, directly impacted the EDCOS Innovation Unit since it is made up of complicated and interrelated steps such as, for instance, defining a business strategy, dismissing or hiring new resources, redirecting funds, and requiring permissions and authorizations

from the directory committee. She held a broader perspective on scientific advancement than the one posited by the American philosopher and historian of science, which confirmed that management, financial and organizational issues all play a critical role in the possibilities to determine scientific change (Ankeny and Leonelli, 2016). The complications inherent in restructuring EDCOS IT environments became enmeshed in the research environment sustained by the digital knowledge infrastructures of HfA, by “freeing” the process of collaboration and the actors who inhabited the collaboration space.

In Francesca’s office, we attended a Zoom meeting with the new project manager, and the first thing noticed was how he introduced the meeting to us: “It is going to be boring for you, but since you are the anthropologists observing us—” he said, smiling upon us. Star’s (1999) insight into the dull materialities of infrastructures seemed to lurk in that Zoom moment. All those supposedly tedious tasks were recorded in GitLab space, a site for collaborative software development and management that also serves as a forum for discussion: programmers can access software repositories and review code changes and upgrades. Using English IT jargon, they examined the nominalization of activities listed in GitLab, which they envisaged as an essential step in building the foundational blocks of a future open-data platform. Nominalization can help establish a coherent framework to solve application and network issues, log-in problems, and bugs. They discussed the relevance and implications of each activity, then developed inferences by comparing them to other activities concerning different use cases. They then relabeled each activity and moved on to the following issue. Many application issues regarding the HfA use case followed the same protocol. Francesca’s words foregrounded all she was learning from and bringing to that experience:

I come from academia, so I am used to looking at the broad picture. Looking at details [how to rename activities and tasks] might be boring, but it helps you develop research questions that are broader and concern the philosophy of the platform as environment.

Francesca’s enlightening phrase “the platform as an environment” revealed how digital infrastructures constantly require acts of maintenance and repair. Making space for Francesca’s perspective highlighted the use of the word “environment” as a metaphor for the platform’s coherence and integrity, which surprisingly sharpened and enriched our thinking of “data as environment.” The scaffolding of the platform seemed to progress quickly and successfully when the project manager joined the team, but the implementation of the EM visualization tool never reached the prototype stage. Following Sofia, Francesca and Matteo helped us gain a greater perspectival awareness of HfA’s

slippages and possibilities. Their collaboration reminded us to think of data not as inert units of analysis and information but as nested, emergent environments that cannot be alienated from the social relations of research.

Seeing data as environments through slippages

Thus far, we have underscored the importance of reflecting on the operationalization of digital knowledge-making infrastructures to foster multidimensional notions of health. We have argued that data and digital knowledge infrastructures are embedded in the intersubjective and multimodal sociality of research. To think of data as environment materializes the socio-material and affective dimensions of knowledge generation by giving prominence to how data and the environments opened up are co-constitutive. In the following paragraphs, we discuss slippage as a sensitizing concept and analytical tool to think of “data as environment.”

Prior studies of infrastructure show that, when functioning as designed, many aspects of infrastructure are rendered invisible or “unexciting” for most users, as they fall into the backstage of daily activities and work routines (Harvey et al., 2016; Star, 1999). Only when infrastructures break down does one become aware of their contribution to daily life. Our ethnography adds to this dichotomous rendering of infrastructure (invisible/working or visible/breaking down) an intermediate stage, in which infrastructure works as a flexible and open structure of contact, thanks to—and not despite—its slippages. These are unavoidable infrastructural conditions determined by the co-existence of different uses of data and the tensions that arise in dealing with the relational openness and multiplicity of data as environments. Slippages result from the co-presence of heterogeneous disciplines working in conjunction with various data uses, affective dispositions and interpretative contexts, demanding new routines that exceed the computational realm. Two types of slippages emerge from our ethnographic illustration: slippages of access and co-labor.

The first type of slippage refers to issues of access: how to search within and isolate data sources, how to inspect and scrutinize different open-data formats, how to sort through open data structured according to regional and national data catalogues’ standards, how to choose the correct format for data reuse in compliance with a specific data-management pipeline, and how to engage with potential users and re-users of the prototype of the visualization tool. As mentioned before, access matters for both use and value. An open, accessible database could help feed science, especially in a health emergency when information becomes valuable for scientific productivity and speed. Access, however, opens up just one level of IT environments. Data’s technical transparency does not correspond with

researchers’ capacity to understand data and learn how to reuse relevant information derived from open databases. Datasets must also be “exposed,” that is, made reusable through a process of data enrichment by integrating meta-data, such as the DOI, and by using specific semantic standards. At stake is both the ability of users to reuse the dataset and extract information from it, as well as the ability of the data provider to expose the dataset to the community in an appropriate manner. We are not simply describing, here, an interdisciplinary space in which researchers feel confused because they do not have the technical knowledge base to understand the intrinsic qualities of data.

Slippages of access can take a negative cast when they demand a learning process that takes time and energy and requires a form of “affective regulation” to ensure that the collaboration does not fall apart (Callard and Fitzgerald, 2015). At the same time, however, slippages of access can open up experimentation practices concerning what can be done with the data available, emphasizing the potential to unhinge disciplinary rigidities, hierarchies and knowledge constructions. Slippages of access broaden the technical context of choices and valuation, implying an active implication in the intersubjective context of research, a form of self-reflexivity. This self-reflexivity brings us to a second type of slippages: those linked to co-labor’s temporalities and digital socialities.

This type of slippage refers to the uses of data being located within multiple timeframes and digital affective socialities. Participation in data-driven collaboration requires a temporal perspective to illuminate the uncertain, non-linear temporalities that permeate the crafting of scientific knowledge: the time to build mutual peer-to-peer trust in research relationships; the time to decide how and where to allocate time and energy to conduct specific tasks; and the disciplinary temporalities for the co-production, co-formalization and cross-validation of knowledge. While interdisciplinary collaborations are well-known opportunities to foster scientific reputation and advance scientific careers, what Sofia and Matteo’s experiences put in a different light is the temporalization of learning and understanding across disciplines. Sofia and Matteo help us understand that temporalities of social relationships, project time, time of emergency, career time and disciplinary time interweave in ways that they cannot anticipate. In addition, in the socio-political conjuncture of COVID-19, the effects of slippages determine a temporal tension between the accelerated temporalities of pandemic technoscience (i.e. the urgency to find social and technical fixes to a multifaceted health crisis) and the temporalities of collaborative science and knowledge co-production. HfA demonstrates that slippages can slow down scholarly communication; delay the achievement of a task or goal, such as drafting and revising for the final submission of a research article; or misunderstand the content of a research communication.

To explicate the temporality of slippages, we need to reconnect digital knowledge infrastructures to their generative processes. We need to bring data into a conceptual unity with the formation of knowledge we have about them, which is an outcome of an intersubjective, multimodal process that becomes constitutive of data.

Slippages of access and co-labor appear in the form of intermittent visibility. They manifest as “gatherings of concerns” that, in our case, do not necessarily receive direct intervention (de la Bellacasa, 2017: 18). Participation in digital knowledge infrastructures implies juggling multiple tasks and the clashes that originate while moving across online and offline environments of data use. Data as environments are not a fixed condition but arise from the contextual capacity of a researcher’s self-reflexivity. These slippages make data as an environment acknowledgeable, something that we can observe. We propose to see data as crafted through the dense interrelationships propelled by data use and their excessive comprehensiveness. Such interrelationships, frequently, do not unidirectionally define and causally entangle how researchers, data processing technologies, biosocial components, mediated territories, statistical institutions and research organizations shape one another. In addition, data are lived, shaped, and made comprehensible, detestable, and lovable through uses and slippages. It might be difficult today to isolate data sources and integrate them to observe various phenomena for scientists with varying computing experience. Awareness of these intermittently visible dynamics is needed to activate the perception of responsibility toward what scientific researchers are studying and stimulate the desire to listen, observe, and understand what is at stake with the data-driven knowledge creation they become part of.

Surprisingly, this environmental framing of data leads to a distinct ontological and epistemological positioning of digital knowledge infrastructures in pandemic technoscience. Although not intentionally, Francesca, Sofia, and Matteo broadened the context of the infrastructure and raised questions about the interrelations between HfA and the evolving socio-political context in which digital knowledge infrastructures operate. From this perspective, “data as environment” interweaves the local dimensions of HfA—that is, the technical, logistic and social aspects—with the planetary aspiration of digital knowledge infrastructures to safeguard health. This finding reflects that of Star and Ruhleder (1994: 120), who also found that what appear as local, concrete issues “broaden the context of choice and valuation” as they raise questions that are “trans-contextual,” questions that elude local infrastructural problems. The specificity of HfA further emphasizes how use determines data’s multiple ontological configurations and related affective contours. Sometimes data are considered perfectible “raw” material; sometimes, they become a currency of scientific collaboration and reward; in other instances, they appear as an implicit, commodified trade

object that does not necessarily lead to a knowledge disclosure. Data result from intersubjective and multimodal relationships. Before being remodeled according to specific workflow management protocols, whenever data are embedded in a knowledge production process, the intersubjective, affective and multimodal approach is reactivated in ways that continuously reconfigure data’s status as commodified information units.

Furthermore, data use becomes connected to the concept of their maintenance, in what we interpret as the process of “learning to be attentive”: asking what is at stake, paying attention and crafting responses (van Dooren et al., 2016). What we describe, here, is a mode of attention and not yet one of practical intervention. This points to if and how scientists can draw and adjust connections with slippages. Data as environments are shaped and lived through the uses, encounters and slippages between knowledge systems, knowledge-processing tools and knowledge subjects. Maintenance offers support and inspiration for co-labor when researchers can draw connections between the research’s IT and socio-political environments opened to view by data. Digital knowledge infrastructures, when they provide a stable and coordinated but flexible framework to conduct research activities, create the conditions that allow researchers to map, understand and manage connections and slippages. Maintenance may help researchers orient, “read” the complexity of the datascape, and bring forth fundamental questions about whether research projects enable scientific subjects with particular desires, needs, aspirations, and imaginaries to formulate questions that take slippages as matters of importance in OH systems of knowledge. In some ways, consistent with the critical literature on OH, this research shows that the steering power of the OH vision and the “technological sublime” of Big Data become dashed—but at the same time, potentially possible—through the difficulties of their realization (Hinchliffe et al., 2021: e231).

Maintenance points to how working with data asks for the creation of a bridge language. Researchers need to learn the language of data scientists. Data scientists need to know how technologies apply across different research contexts of data use, because what works for one group does not necessarily work for another. Although Francesca, Sofia and Matteo benefited from each other’s expertise, they still needed to co-learn some degree of each additional expertise. In addition, Francesca, together with her team of developers, had to acquire the expertise and vocabulary to understand and implement Matteo and Sofia’s required functionalities. To borrow Hinchliffe and colleagues’ (2021: e231) words, “the crucial question is therefore not how digital technology might deliver public health [*one health*] but how a healthy public [*one healthy public of scientists*] might arise from new forms of communication and knowledge generation.” Thus far, the opening of human health to the planet as a whole is rooted in the

social process of the co-production of knowledge. The achievement of the OH vision is determined by how working with technologies for collecting and processing data allows researchers to activate the perception of responsibility toward what we are studying; to stimulate the desire to listen, observe, and understand; and, as a result, to ask fundamental questions and call for answers capable of making sense of what is happening.

In seeking to underscore the inescapable and ambivalent constitution of infrastructure through an environment that includes slippages, we want to interpret the lifecycle of HfA as more than an infrastructural failure. Ethnographically studying and historicizing HfA collaborative data practices have allowed us to see that slippages are embedded in a social and historical context of lengthened, comprehensive transformations regarding interdisciplinary science in a data-driven world. Slippages are not simply constraints to scientific advancement; they also offer an opportunity for researchers “to meet half way” (Barad, 2007) and draw connections among socio-technical, affective, intersubjective, technological, political, and epistemic dimensions. If data is more than information technology, if data is an environment, implementing the OH vision should consider digital technoscientific socialities. There is a need to recognize and align the different systems of knowledge that guide how researchers define what is valuable for applying multidimensional notions of health as OH. To fill in the gaps created by inefficient communication, to draw connections among different uses of data, to mend the knowledge holes of open-data technologies, to learn how and where to hold attention to things, to attune to a variety of affective conditions, and to build a shared valuation framework for research activities: these all are acts of maintenance determined by the willingness to connect environments opened to view by data. Maintenance, thus, scales up openness—not merely including data but “data as environment”—bringing us closer to the social and political context in which digital infrastructures of collaborative science are embedded.

Conclusion

This article has reflected on the notion of “data as environment” in relation to the emergence and operationalization of multidimensional approaches to health. In particular, it examined the advancement of collaborative data-driven research within the context of the COVID-19 emergency, when anxieties to understand an evolving pandemic revamped views of science as a pursuit of public knowledge and technoscientific development, albeit reinforced by Big Data mythology (Beaulieu and Leonelli, 2022; Leonelli, 2021).

There are many ways through which data as environments unfold, and our case study is but one example. We

do not claim to have illustrated all possibilities, but our ethnographic research highlights how data, far from being enclosed units of information, are a structure of contact that puts into relationship—both empirically and theoretically—knowledge subjects, knowledge systems and knowledge-processing tools.

Unsurprisingly, openness doesn’t apply to every layer of the environment opened up by data. Slippages show that openness is not only a technical matter. To advance scientific concepts, such as OH, through collaborative data-sharing initiatives depends not only on the ability to create interoperable units of information, user-friendly interfaces, platforms and tools. It also requires the ability to infuse reflexive thinking across disciplines and computing expertise through which researchers create and maintain the socio-material conditions for the co-production, validation and dissemination of knowledge.

If data-driven collaborative practices are still difficult to grasp, that is because they are the result of an emergent and multifaceted convergence of socio-technical, cultural and technological aspects. If we conceive data as a structure of contact activated by use and maintenance—that is, an environment—digital infrastructures of collaboration should not be conceived as passive entities that can be acted upon, manipulated and controlled once they are in place. Digital knowledge infrastructures are not merely socio-technical apparatuses at the service of researchers. They are built and legitimated in the intersubjective process of research. From this perspective, if the stated mission of OH is to include the environment in the understanding of health, acknowledging that the OH vision is nowadays achieved through data-driven science helps us see that the concept of environment applies not only to health but to data, too, indeed, the same data that promise to make OH operable. In other words, considering environments as data invites us to think of data as environments too.

Acknowledgements

This article is part of a project that has received funding from the European Union’s Horizon 2020 Research and Innovation Programme (GA n. 949742 ERC-HealthXCross). The article has benefited from insightful feedback from Sabina Leonelli and the Egenis scientific community during Barchetta’s visit. We also thank Tone Walford for her constructive feedback on an initial draft and Hannah Landecker and Joe Dumit for their engaging discussions, which have inspired the scaffolding and conceptual framework of this paper. Additionally, we express gratitude to HfA’s PI and participants for their collaboration and openness to engage.

Declaration of conflicting interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the European Research Council 2020 – Starting Grant (grant number GA n. 949742 ERC-HealthXCross).

ORCID iDs

Lucilla Barchetta  <https://orcid.org/0000-0001-5110-3227>

Roberta Raffaetà  <https://orcid.org/0000-0002-0680-1613>

Notes

1. This article/book is part of a project that has received funding from the European Union's Horizon 2020 Research and Innovation Programme (GA n. 949742 ERC-HealthXCross).
2. For an anthropological discussion of the concept of “excess,” see: de la Cadena (2015) and de la Cadena and Escobar (2023).
3. In November 2021, when the field research started, the total number of HfA members was around 40 but it decreased significantly over time.
4. Barchetta Lucilla conducted most of the ethnography, wrote sections 2, 3 and 4 and contributed to the writing of the Introduction and Conclusion. Raffaetà Roberta gained access to the field and conducted the first months of fieldwork, wrote the Introduction and Conclusion and collaborated to the other sections.)
5. <https://ourworldindata.org/excess-mortality-covid#how-is-excess-mortality-measured> (accessed on 15 February 2023).
6. “That is, the agential cut enacts a resolution within the phenomenon of the inherent ontological (and semantic) indeterminacy. In other words, *relata* do not preexist relations; rather, *relata-within-phenomena* emerge through specific intra-actions. Crucially, then, intra-actions enact *agential separability*—the condition of *exteriority-within-phenomena*.” (Barad, 2007: 140, emphasis in original)

References

- Ahmed S (2019) *What's the Use?: On the Uses of Use*. Durham, NC: Duke University Press.
- Ankeny RA and Leonelli S (2016) Repertoires: A post-Kuhnian perspective on scientific change and collaborative research. *Studies in History and Philosophy of Science* 60: 18–28.
- Baedke J and Buklijas T (2023) Where organisms meet the environment: Introduction to the special issue ‘What counts as environment in biology and medicine: Historical, philosophical and sociological perspectives’. *Studies in History and Philosophy of Science* 99: A4–A9.
- Barad K (2007) *Meeting the Universe Halfway. Quantum Physics and the Entanglement of Matter and Meaning*. Durham & London: Duke University Press. Available at: https://drive.google.com/file/d/1WirhgZyXw7dJDLTa0yF_Vl6wpiH2e-k/view?usp=drive_link.
- Beaulieu A and Leonelli S (2022) *Data and Society. A Critical Introduction*. Sociology. London: Sage Publications Ltd. Available at: <https://uk.sagepub.com/en-gb/eur/data-and-society/book269709> (accessed 29 August 2022).
- Benis A, Tamburis O, Chronaki C, et al. (2021) One Digital Health: A unified framework for future health ecosystems. *Journal of Medical Internet Research* 23(2): e22189.
- Borgman CL (2012) The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology* 63(6): 1059–1078.
- Bowker GC (2000) Biodiversity datadiversity. *Social Studies of Science* 30(5): 643–683.
- Callard F and Fitzgerald D (2015) *Rethinking Interdisciplinarity across the Social Sciences and Neurosciences*. Basingstoke: Palgrave Macmillan. Available at: <http://dx.doi.org/10.1057/9781137407962> (accessed 9 June 2023).
- Craddock S and Hinchliffe S (2015) One world, one health? Social science engagements with the one health agenda. *Social Science & Medicine* 129: –4.
- de Koning M, Meyer B, Moors A, et al. (2019) Guidelines for anthropological research: Data management, ethics, and integrity. *Ethnography* 20(2): 170–174.
- de la Bellacasa MP (2017) *Matters of Care. Speculative Ethics in a More Than Human World. Posthumanities*. Minneapolis and London: Minnesota University Press. Available at: <https://www.upress.umn.edu/book-division/books/matters-of-care> (accessed 2 March 2023).
- de la Cadena M (2015) *Earth Beings: Ecologies of Practice across Andean Worlds*. Durham, NC: Duke University Press. Available at: https://books.google.it/books?id=JN_orQEACAAJ.
- de la Cadena M and Escobar A (2023) Notes on excess: Towards pluriversal design. In: *Design for More-Than-Human Futures*. New York: Routledge, 29–50.
- Delfanti A (2013) *Biohackers: The Politics of Open Science*. London: Pluto Press.
- Dilger H, Pels P and Sleeboom-Faulkner M (2019) Guidelines for data management and scientific integrity in ethnography. *Ethnography* 20(1): –7.
- Douglas-Jones R, Walford A and Seaver N (2021) Introduction: Towards an anthropology of data. *Journal of the Royal Anthropological Institute* 27(S1): 9–25.
- Dumit J (2021) Substance as method (shaking up your practice). In: *Reactivating Elements. Chemistry, Ecology, Practice*. Duke University Press, 175–185. DOI: 10.1515/9781478091691-016.
- Edwards PN (2017) Knowledge infrastructures for the Anthropocene. *The Anthropocene Review* 4(1): 34–43.
- Edwards PN, Jackson SJ, Chalmers MK, et al. (2013) *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. 1 May. Ann Arbor: Deep Blue. Available at: <https://escholarship.org/uc/item/2mt6j2mh> (accessed 22 November 2021).
- Farman A and Rottenburg R (2019) Measures of future health, from the nonhuman to the planetary. *Medicine Anthropology Theory* 6(3): 1–29.
- Gabrys J (2016) *Program Earth. Environmental Sensing Technology and the Making of a Computational Planet*. Minneapolis and London: University of Minnesota Press. Available at: <https://www.upress.umn.edu/book-division/books/program-earth> (accessed 3 October 2022).
- Gibbs EPJ (2014) The evolution of One Health: A decade of progress and challenges for the future. *Veterinary Record* 174(4): 85–91.
- Gitelman L (ed) (2013) *Raw Data Is an Oxymoron*. Cambridge, MA and London: MIT Press.

- Haraway D (1987) A manifesto for cyborgs: Science, technology, and socialist feminism in the 1980s. *Australian Feminist Studies* 2(4): 1–42.
- Harvey P, Jensen C and Morita A (2016) *Infrastructures and Social Complexity: A Companion*. London: Routledge. DOI: 10.4324/9781315622880.
- Henare A, Holbraad M and Wastell S (2007) *Thinking Through Things: Theorising Artefacts Ethnographically*. London: Routledge.
- Hinchliffe S, Manderson L and Moore M (2021) Planetary healthy publics after COVID-19. *The Lancet Planetary Health* 5(4) Elsevier: e230–e236.
- Horst HA and Miller D (eds) (2013) *Digital Anthropology*. London: Routledge.
- Keusch GT, Amuasi JH, Anderson DE, et al. (2022) Pandemic origins and a One Health approach to preparedness and prevention: Solutions based on SARS-CoV-2 and other RNA viruses. *Proceedings of the National Academy of Sciences of the United States of America* 119(42): e2202871119.
- Kitchin R (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. Los Angeles: Sage.
- Leboeuf A (2011) *Making Sense of One Health: Cooperating at the Human-Animal-Ecosystem Health Interface*. Health and Environment Reports 7. Available at: <https://www.ifri.org/fr/publications/etudes-de-lifri/making-sense-one-health-cooperating-human-animal-ecosystem-health> (accessed 29 September 2022).
- Leonelli S (2016) *Data-Centric Biology: A Philosophical Study*. Chicago: University of Chicago Press.
- Leonelli S (2021) Data science in times of pan(dem)ic. *Harvard Data Science Review* 3(1): 2–30.
- Leonelli S (2023) *The Philosophy of Open Science*. Cambridge: Cambridge University Press.
- Leonelli S and Tempini N (eds) (2020) *Data Journeys in the Sciences*. Cham: Springer International Publishing. DOI: 10.1007/978-3-030-37177-7.
- Meloni M (2014) Biology without biologism: Social theory in a postgenomic age. *Sociology* 48(4): 731–746.
- Niewöhner J and Lock M (2018) Situating local biologies: Anthropological perspectives on environment/human entanglements. *BioSocieties* 13: 681–697.
- Okune A, Hillyer R, Albormoz D, et al. (2019) Whose infrastructure? Towards inclusive and collaborative knowledge infrastructures in open science. In: *Connecting the Knowledge Commons: From Projects to Sustainable Infrastructure: The 22nd International Conference on Electronic Publishing – Revised Selected Papers*. Marseille: Open Edition. Available at: <https://hal.science/hal-01816808> (accessed 20 January 2023).
- One Health High-Level Expert Panel, Adisasmito WB, Almuhairi S, et al. (2022) One Health: A new definition for a sustainable and healthy future. *PLOS Pathogens* 18(6): 1–4.
- Pedersen MA (2023) Editorial introduction: Towards a machinic anthropology. *Big Data & Society* 10(1) SAGE Publications Ltd: 20539517231153804.
- Pels P, Boog I, Florusbosch JH, et al. (2018) Data management in anthropology: The next phase in ethics governance? *Social Anthropology* 26(3): 391–413.
- Pink S, Sumartojo S, Lupton D, et al. (2017) Mundane data: The routines, contingencies and accomplishments of digital living. *Big Data & Society* 4(1). DOI: 10.1177/2053951717700924.
- Poirier L and Costelloe-Kuehn B (2019) Data sharing at scale: A heuristic for affirming data cultures. *Data Science Journal* 18(1). DOI: 10.5334/dsj-2019-048.
- Ribes D, 2006. *Universal informatics: Building cyberinfrastructure, interoperating the geosciences*. PhD diss. UC San Diego. Available at: <https://escholarship.org/uc/item/2gf7b45n> (accessed 26 July 2023).
- Richardson SS and Stevens H (eds) (2015) *Postgenomics: Perspectives on Biology after the Genome*. Durham, NC: Duke University Press.
- Sharma K (2015) *Interdependence: Biology and Beyond. Meaning Systems*. New York: Fordham University Press.
- Star SL (1999) The ethnography of infrastructure. *American Behavioral Scientists* 43: 377–391.
- Star SL and Ruhleder K (1994) Steps towards an ecology of infrastructure: Complex problems in design and access for large-scale collaborative systems. In: Proceedings of the 1994 ACM conference on computer supported cooperative work, New York, 22 October 1994, pp. 253–264: CSCW. Association for Computing Machinery. DOI: 10.1145/192844.193021.
- van Dooren T, Kirksey E and Münster U (2016) Multispecies studies: Cultivating arts of attentiveness. *Environmental Humanities* 8(1): 1–23.
- Verhulst SG (2015) Data collaboratives: Exchanging data to improve people’s lives. In: *Medium*. Available at: <https://sverhulst.medium.com/data-collaboratives-exchanging-data-to-improve-people-s-lives-d0fcfc1bdd9a> (accessed 3 October 2022).
- Woods A, Bresalier M, Cassidy A, et al. (2018) *Animals and the Shaping of Modern Medicine: One Health and Its Histories*. Cham: Springer International Publishing.