

Mike Fortun and Kim Fortun. In press. "What's so funny 'bout PECE, TAF, and data sharing?" *Collaborative Anthropology Today: A Collection of Exceptions*. Dominic Boyer and George E. Marcus (eds.) Cornell University Press.

What's so funny 'bout PECE, TAF, and data sharing?

Mike Fortun and Kim Fortun
with

Vinay Baidur, Brian Callahan, Jessica Chandras, Brandon Costelloe-Kuehn, Thomas Depree, Sam Elrahman, Erica Fletcher, Govind Gopakumar, Rodolfo Hernandez, Jason Baird Jackson, Madhura Joglekar, Scott Kellogg, Ali Kenner, Aalok Khandekar, Jobby Kunjachen, John Mathew, Maria Michails, Alli Morgan, Lindsay Poirier, Dan Price, Deepa Reddy, Rohit Negi, Surajit Sarkar, Pankaj Sekhsaria, Prerna Srighyan, Sharon Traweek, Vivian Wong, et al. (AKA The World-PECE.org Network of Networks)

Here we discuss four different ways we are involved in and with collaborative projects. They share in many ways a form, shape, or style, and may be imagined as nested within each other, like *matryoshka* dolls -- if *matryoshka* could be rendered hyperdimensional, so that there were no smaller and larger "scales" but each could suddenly shift places and relationships. Each of our collaborative projects is in conversation with, informed by, or at least rubs up against each of the others, sometimes the framing form, at other times the framed content. Their differences are less a matter of scale (smaller to larger collaborations) or temporal development (early to later collaborations), and more a matter of any given collaborative project being both inside and outside to another, a flickering switch of figure and ground, an exchange in which one form becomes content for another, whose content in turn (in)forms the next, setting in motion chains of demands and possibilities that animate the collaborations and set the stage for new lines of experiment, growth, and change.



The first one we discuss (but not logically, organizationally, or historically first) is PECE, the Platform for Experimental Collaborative Ethnography, the digital infrastructure we have been developing for several years to support new collaborative projects in anthropology. The next collaborative project presented here is the much longer-standing collaboration of The Asthma Files (TAF), an experimental ethnographic research project that eventually led to the conceptualization and development of PECE. TAF is now in large part realized through our 6+ Cities Research project on air quality science and governance in comparative cultural context, the third project discussed here. All of those can be thought of as taking shape within another looser, larger, more dispersed, and more sporadically collaborative (for us) layer, the Digital Practices in History and Ethnography Interest Group (DPHE-IG) we organized within the Research Data Alliance (RDA), a global collaboration of individuals and institutions working to make data more easily and openly shareable.

We present each of these through a primarily descriptive style, for several reasons. The first iteration of this essay was written to share with participants in the workshop on collaboration held at the Center for Ethnography at UCI, and for this assignment, an empirical dimension was explicitly troped as more important than the theoretical. As George Marcus wrote to us in an email: "It is fine if you

would like to include broader conceptual or methodological reflections on collaboration, but what is most important is that we get a detailed sense of what your collaboration has involved so far so that we are all familiar with each other projects by the time we gather.” But beyond that, we also tend to be somewhat “matter-of-fact” about collaboration, collaborative projects, and “the collaborative form.” We of course have our reasons for and commitments to collaboration, many of which we learned from or had reinforced by our fieldwork experiences among scientists and engineers, for whom interdisciplinary collaboration is definitely “a thing”--a thing difficult to define and delineate, let alone actually pull off, but also a much desired and culturally supported “thing.” But collaboration is also in some ways just something that we do. We recognize that it’s not for everyone, and that it’s not right for every project; it has its virtues, but those do not make collaboration any more virtuous than other ethnographic forms; collaboration is vital to future ethnography, but it’s hardly the only way to do that futuring. Our commitments to collaboration are best described as thoroughly experimental: *trying* collaboration is crucial, to build and multiply collaborative projects, so as to understand better what they can and cannot do.

In that sense, the “collaborative form” for us is also the experimental form analyzed by Hans-Jorg Rheinberger as essential to a modern scientific style, in which the limits (of knowledge as a system of signs, of laboratory equipment as a system of material devices) are the active site of a double movement: work *and* play, closure *and* openness, structure *and* de-structure, reliable technical reproduction of phenomena *and* the generation of unanticipated, surprising events and objects. These double movements of experimental systems are, for Rheinberger, how sciences grow and even progress in the sense of becoming more encompassing, more productive, more competent in more situations. Each of our projects is experimental in this sense as we’ve learned it through Rheinberger’s analysis of scientific work and change, and not simply in the sense of doing something new and avant garde: we experiment by working the structures and infrastructures (technical, organizational, interpersonal) that provide the stable grounds for producing, saving, and sharing our individual and collective data, and by playing the inevitable movements, insufficiencies, and open edges of these (infras)structures to keep things lively, interesting, and new.

In keeping with these dual “matter-of-fact” and experimental commitments, in which what we call “light structure” enables both stability and change, and in which dazzling theorization is at least partly deferred as one pursues more mundane outcomes, we decided to adopt for this essay a well-known and productive “light structure,” a rhetorical convention or prompt which formed the basis of the journalists’ “Five W” questions--in an older version, the “seven circumstances” ascribed to Hermagoras of Temnos: *who, what, when, where, why, in what way, by what means* (in Latin, *quis, quid, quando, ubi, cur, quem ad modum, quibus adminiculis*)?¹ As we elaborate below, in the PECE platform and in the collaborative projects it supports we rely extensively on such “light structures:” minimal, open-ended forms that, when instantiated in different technologies or media, help facilitate collaboration and comparative analysis. Responding to them is always an interpretive act (*what does “when is the collaboration?” ask for?*), but one whose differences are always in contact with shared continuities and connections.

PECE: Platform for Comparative Experimental Ethnography

¹ We cite the collaborative entry at https://en.wikipedia.org/wiki/Five_Ws. Wikipedia notes the many variations of this device (now widely regarded as “old-fashioned”), including Rudyard Kipling’s rendition in “The Elephant’s Child” from *The Just-So Stories*:

*I keep six honest serving-men
(They taught me all I knew);
Their names are What and Why and When
And How and Where and Who.*

Who: A small collaborative “Design Team” first physically centered at Rensselaer Polytechnic Institute (now more geographically dispersed) has been at the core of PECE development for the past several years (<http://pece.readthedocs.io/en/latest/team.html>): Mike Fortun and Kim Fortun initiated and lead the project, which has depended on (collaborations depend on dependencies) two of their graduate students in the Department of Science and Technology Studies who are also highly technically skilled, Lindsay Poirier (who has earned the title “Lead Platform Architect”) and Brian Callahan (our Lead Open Knowledge Developer). Both Lindsay and Brian have utilized their work with PECE and their engagement with digital/technical worlds more generally as part of their fieldwork for different, broader PhD dissertation projects. Other current and former graduate students at RPI also contributed to PECE’s design and implementation, including Alli Morgan, Ali Kenner, Brandon Costelloe-Kuehn, and Erik Bigras. Rensselaer computer science PhD student Dominic DiFranzo and UCLA anthropology graduate student Luis Felipe Murillo also circulated through our group for a period of time and made important contributions to the platform. After a brief (failed) experiment with a Plone developer (Plone is a content management system, or CMS, that we explored in an earlier collaboration with Dan Price at the University of Houston), the actual coding/building of PECE was accomplished in a Drupal by contracting in 2015 with Taller, a Brazilian “digital business studio that turns ideas into innovative business” (<http://taller.net.br/en/#section-members>); Renato Vacsoncellos Gomes was our main developer at Taller, and he has continued to work independently (and enthusiastically!) with us since then. Another Boston-based Drupal development company, the “worker-owned cooperative” Agaric (<http://agaric.com/>), has made more limited contributions.

What: From our most recent (collaboratively produced) Project Statement:

The Platform for Experimental, Collaborative Ethnography: History and Current Capabilities

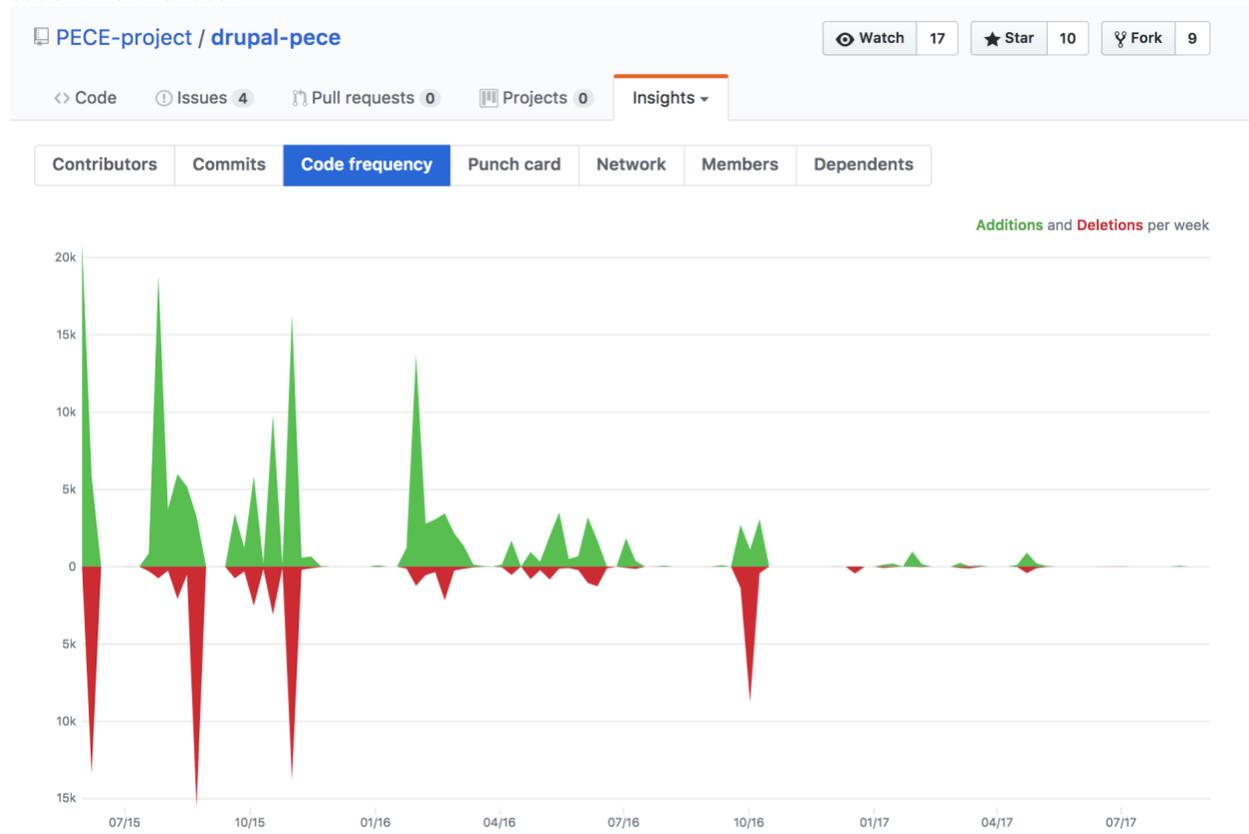
PECE (worldpece.org) is an open-source, Drupal-based platform designed to support a wide range of collaborative humanities projects. PECE provides a space for geographically dispersed researchers to share primary materials (such as field notes, grey matter, photographs, and recorded interviews), provides tools supporting analysis and interpretation of these materials, and allows researchers to experiment with new ways of publishing their results, addressing diverse audiences. A core goal is to support greater collaboration among empirical humanities researchers, and between these researchers and researchers in other fields. The project is working to theorize, methodologically enact, and technically support “collaborative hermeneutics.”

PECE is shaped by the empirical demands and theoretical tenets of experimental ethnography, but is designed to be used by an array of research groups, helping build out a rich ecology of interoperable digital projects that link researchers in different fields, and with diverse stakeholders. The platform design emerges from concrete ethnographic practices and (newly invented) collaborative work flows, inflected by poststructural understandings of language, meaning and knowledge. PECE’s “design logics” are meant to stay on the surface, open to debate and revision. This effort is sustained by the PECE design group, on PECE’s own platform, and platform users are encouraged to continually help develop them.

A signature feature of PECE is the way it supports the archiving, sharing, and collaborative use of a diverse (and ever growing) set of “structured analytics”--sets of questions researchers use to identify, generate, and interpret empirical humanities data. Archiving and sharing these structured analytics exposes a work flow and intellectual process often invisible and tacit in humanities research. Digital research infrastructure in the natural and social sciences often exposes such work flows, aiming for increased reproducibility of research results; digital humanists face a related but somewhat different challenge and goal. In the humanities,

fine-grained exposure of workflows and intellectual processes can convey the depth and rigor of humanist analysis--what John Dewey would have called not its reproducibility, but its warrantability--while also supporting new forms of peer review and collaboration early in the research process. With capacity to share “structured analytics,” researchers can 1) better examine why and how data was created, 2) analyze data using multiple frameworks, from diverse traditions of thought (open to juxtaposition), and 3) analyze data collaboratively, leveraging robust interpretive pluralism among researchers rather than simple reproducibility. This is supported with a tool that allows researchers to integrate multiple annotations produced by different researchers using PECE’s shared structured analytics, allowing easy visualization of multiple researchers’ analyses and interpretations at all stage of the research process.

When: For the better part of 2015-16, the PECE Design Team had weekly two-hour face-to-face meetings, with regular email and skype contact happening around that. Even if that tempo has since abated, PECE is, out of all of our collaborations, the one that is the most constant even if not the most active presence in our attention-scape. For most of 2016, Lindsay always had a skype chat window open with Taller during each of our six “legs” of development; you can see the record of code additions and deletions in the graph below. Lindsay was the central figure in the techno-social work of this collaboration (again, earning her the title “Lead Platform Architect” in many of our documents); between cajoling Taller to keep to schedule, testing new features as they were developed and either approving them or identifying bugs, and writing the “user stories” that would define the next leg of development work, she logged the most hours of any of us. And she continues, as Lead Platform Architect, to supervise all new software development, still keeping a Skype window onto Brazil to stay in touch with Renato.



Platform development never went as quickly as we or our collaborators wanted, and we consider PECE to be in “permanent beta” mode: it works well, but it could always work better, and we are always planning and writing grants for that. The constant drone of software development was punctuated by other events in time; “milestones” like the first public release of PECE (May 2016) on GitHub became important markers that quickened the collaboration’s pace.

The best short answer, then, to the light structure question “when is PECE?” is: more or less always.

Where: Aside from the Brazilian Drupal developers, this collaboration ran almost entirely at RPI among its faculty and graduate students. Members of Rensselaer’s “Tetherless World” constellation of computer, web, and data scientists supported us with occasional feedback and advice, web hosting, and a bit of funding. It’s also where we convened several design and user-feedback workshops where we involved other scholars with professional or research interests and commitments to digital developments, like Jason Baird Jackson and Sharon Traweek.

Another answer to the “where?” analytic is: GitHub. The PECE collaboration resulted in writing code into GitHub where, in the time of permanent beta, it quietly awaits further “forks” that will develop it anew. And in a slightly different instantiation, PECE’s current code is zipped up into a “distro,” for anyone to download, install on a server, and open into a new project.

Why: PECE is part of what we’ve thought of as “an infrastructural moment in the human sciences” (Fortun and Fortun 2015). A differential continuation of anthropology’s “experimental moment” of the 1980s described by Marcus and Fischer (1986), we described the 2000s as an “infrastructural moment” characterized by new (digital) technological demands and possibilities, with issues of open access publication central to them. If ethnography was to continue to be experimental and grow to be more collaborative, those collaborative experiments need digital infrastructure designed (and not simply adopted from elsewhere) to support them. We became great admirers of recent infrastructure projects in cultural anthropology with similar intent, like openfolklore.org, digitalhimalaya.com, and mukurtu.org. There may be many collaborative projects in anthropology, but only a few of those have tried to design and provide new digital scaffolding to multiply the attempts.

PECE can also be thought of as an experiment in collective curation, one that provides the digital infrastructure to distribute collaboratively the work of compiling, documenting, characterizing, archiving, interpreting, and making public an agglomeration of “ethnographic data” bearing on complex, heterogeneous, cultural phenomena. Why? Not because we were fueled by some dream of exhaustive totality (Lemov 2015), nor because we thought every ethnographer needed to make all of their data about everything and everyone completely open and accessible. But there are many ethnographers, we reckoned, in projects like ours where issues of privacy and research ethics were manageable (we “study up,” for the most part), and simply making *more* of our newly created or otherwise singular data *more* accessible to *more* researchers would just be good, or at least a good experiment to try, even though it would not lead us all into a state of *anthroparousia*. Just as “raw data” is at least somewhat oxymoronic (Gitelman 2013), so too is “private data:” the more public data can be made to be, the more it can be data. We wanted PECE to provide a means, in a limited number of cases, to render into “data” that which might otherwise remain as ephemera, “grey matter,” or just one of the dozens of interviews conducted every day, surely, by an ethnographer somewhere that would end up—either for reasons of “research ethics,” or the outdated expectations of funding agencies, or a culture of proprietary individualized scholarship—locked away in a file drawer.

In what way: Although the mantra of “more hack, less yack” was never a particularly productive or apt one in the field of digital humanities writ large, the two terms nevertheless sum up much of the

collaborative labor and its division: Kim and Mike worked almost exclusively through talk, while Lindsay, Brian, and others did the typing/coding/hacking--but much yacking, too. Yacking includes reading: we collaboratively read, analyzed, and discussed a wide and varied scholarly literature in digital humanities (see e.g. Gold 2012), especially those dealing with the archival form (Derrida 1998; Brown 2006; Klein 2013; Turin 2011).

By what means: While yacking has been a more or less ongoing feature of the collaboration, the hack came in waves, as funding allowed (again, see the graph above). The development of the data model and “practical policies” (how permissions are handled on the platform, backups, data expiration, etc. etc.) was done with RDA funding (see below), which also paid for some of the Drupal development work. Funding for all the rest of the Drupal programming was pieced together through the NSF grant that supported the 6+ Cities Project (see below), and two RPI internal program grants. Brian mostly gifted his time, knowledge, and skills to PECE; Lindsay has been supported through a research assistantship not through the STS department at RPI but through the Tetherless World Constellation (information, web, and computer science). Consistent with the dominant cultural logics of collaborative labor, each of them has committed more to PECE than what they have been remunerated for.

The Asthma Files (TAF)

Who: This collaboration started as collaborative research project between Kim and Mike Fortun, stemming from their research as anthropologists of the environmental sciences and genomics, respectively, responding to new initiatives in those fields involving asthma as a place where new methods for collaborative “gene-environment interaction” research could be worked out. At the start it involved the Fortuns and their graduate students, former graduate students, and undergraduate students at both RPI and other universities, working on an array of projects, some more independent, some more intricately entangled, all collaborative at some level. Undergraduates have played a significant role in this collaboration; easing their involvement was one of the explicit motivations for building PECE. TAF is open-ended, designed to accommodate and welcome any short- or long-term contributors, on any existing topic or any new ones generally related to breathing, air, respiratory disorders, and the many sciences thereof. Graduate student in the RPI STS department Alli Morgan has been responsible for much of the topical, substantive, and administrative developments, including organizing RPI undergraduate researchers to contribute to the project. Former graduate student Ali Kenner, whose doctoral dissertation spun off from an earlier (pre-PECE) form of TAF, has provided much content to the current project/platform, both through her own energetic research (Kenner 2018) largely through a rotating cast of undergraduate and graduate students in her classes at Drexel University. We know most of these Drexel collaborators only through their work on the platform. As of this writing in mid-2017, much of the activity at TAF occurs through the 6+ Cities Project described in the next section.

What: From The Asthma Files (<http://theasthmafiles.org>) website:

The Asthma Files (TAF) is a collaborative ethnographic research project designed to advance understanding and efforts to address environmental public health challenges around the world. Focusing on dramatic global incidence of asthma and other respiratory illnesses as a starting point, the project spirals out to address growing concern about the health impact of air pollution and associated need to build scientific, clinical and public health capacity to address environmental determinants of human health. Through ethnographic interviews and analysis of scientific publications, policy debates, and media coverage, the project draws together many different ways of approaching environmental public health, aiming to enhance comparative and collaborative perspective.

A key aim is to develop comparative understanding of different styles of both environmental health research and environmental health governance, in different urban and national settings. The project will result in a theoretically robust, empirically grounded conception of (environmental health) research and governance styles, detailing and categorizing different ways of developing environmental health data, advancing the sciences of environment and health, and directing these toward governance of complex problems. The project thus builds on work in the history and anthropology of science on how “thought styles” shape scientific research, and extends it to sociocultural analysis of “governance styles...”

When: TAF has a longer history than PECE, and for many years was the collaborative research project that eventually gave rise first to the conceptualization and then to the design and building of PECE. TAF has existed on different platforms, beginning in 2001 as a Powerpoint file that we (KF and MF) built beginning as a collaboration of two, where each slide was a “file” presenting a concept, institution, researcher, or other such entity drawn from our research illustrating different “asthma knowledges.” The intent was to privilege difference and epistemological pluralism over the consensus-driven, monopolistic approaches favored in dominant scientific discourses. By 2007 TAF had become a wiki to which our students could make their own contributions, diversifying the topics and expanding the scope of TAF.² We started a Zotero group library in 2009 to store the collaboratively compiled bibliographical references; it currently has 44 members and over 3,000 references. By early 2012 we had begun migrating to a platform built with the Plone CMS, in collaboration with Dan Price and the Texas Learning and Computing Center at the University of Houston. Throughout this history, TAF also had weekly meetings, mostly with its RPI members but often including other participants by skype. Those weekly meetings led eventually to the weekly PECE meetings, and the developments in and on Drupal described above, and the eventual migration of TAF to the PECE platform in 2016. (We can only quickly re-emphasize here the importance that the development in the 2000s of increasingly stable and available digital technologies like Skype, Google Docs, Drupal, and Zotero have had on our ability (and desire) to collaborate across geographic and other differences.)

Where: For the first several years of its development (pre-PECE) TAF was a collaboration entirely within the RPI STS faculty and grad students (and a limited number of undergraduates). As the number of collaborators grew, TAF began to develop a focus on specific “asthmatic spaces,” where the emergent geospatial unit was the city: Houston first, then Knoxville, Tehran, Philadelphia, New York City. “Asthmatic spaces” was one of six thematic “cabinets” into which asthma “files” were organized (the others were “accounting for asthma,” “knowing asthma,” “experiencing asthma,” “caring for asthma,” and “communicating asthma.”) This city-centric orientation eventually led to the even more expansive collaboration of the 6+ Cities Project.

Why: Asthma represents the kind of “complex condition” -- chronic, multicausal, confoundingly variable, researched by numerous different expert communities -- that demands collaborative methods and approaches. That statement is true for the scientists who research asthma, for whom “interdisciplinarity” or “transdisciplinarity” or “team science” are valued and ubiquitous (and poorly defined and elusive) terms, and it is as true for us as ethnographers: an anthropology of asthma cannot not be collaborative. The PECE platform was designed to configure the collaboration among different researchers and their contributions to produce what we called “kaleidoscopic perspective.” The intent again was to privilege or foreground explanatory pluralism and difference by juxtaposing varied materials and analyses in ways that might displace or disrupt the drives toward consensus,

² <http://theasthmafiles.wikispaces.com/The+Asthma+Files+Wiki>

harmonization, and totalization that characterize the dominant styles in not only the natural but the human sciences as well.

In what way: Although almost everything about TAF is collaborative, or depends on collaboration, we highlight two collaborative dimensions here that PECE *infrastructures* (to be read as a verb) into the TAF project as core innovations: 1) the annotation module, built around what we called “collaborative analytics;” and 2) what we at first referred to metonymically as “the asthma file” and now call (to differentiate it from the TAF project and generalize it) “the PECE essay.”

1) The PECE annotation module is an instance of “light structure” that encourages and leverages “explanatory pluralism” and “collaborative hermeneutics:” documents, images, audio and video interviews, and any other “artefact” contributed by a researcher can be collaboratively and interpretively analyzed by *n* other researchers who respond to an open-ended series of shared, open-ended questions. The PECE Project Statement describes what collaborative analytics look like in practice:

In this stage, researchers analyse artefacts, oriented by a set of shared questions. A researcher does not have to respond to every question, but each response becomes its own “object” within PECE, making it available for later comparison and combination with other responses to the same question addressed in the analysis of other artefacts. Researchers can also add questions to the shared set of questions...Many annotations, especially early in the work of a project, are of relevant published material, but any object can be annotated: an image, an interview or interview excerpt, a *Nature* article. The annotations are structured to ease sharing and comparing of “notes,” and to pull analysis back to a project’s “shared questions.” As a researcher writes an annotation, the entry window presents her with a series of these questions, to which all other project collaborators are also responding. A researcher may ignore some questions and write extensively on others, but each response becomes...available in a structured way for recombination with other annotations on other materials.

The TAF site, for example, has these kinds of collaborative analytics structured for “Profiling a Data Policy,” “Reading Digital Infrastructure,” “Profiling an Organization,” and “Annotating a Scientific Study” (to name only a few). The latter presents researchers with questions such as “How was the study funded?,” “Describe how this study has traveled. Has it been used in news reports, cited by health officials, or used as the basis of other studies?,” and “How, if at all, does the study address vulnerable populations?” Any user can add a new question to any collaborative analytic; PECE’s digital infrastructure was designed to “credit” such collaborative contributions of questions, so the system values and credits the authorship of questions and analytics as much as it credits the authorship of data and interpretative analysis.

Indeed, every component of these collaborative analytics was *infrastructured* and designed (this consumed a lot of time and money) to be, in effect, its own “data object,” digitally outfitted with enough appropriate metadata to make it credit-able, archivable, searchable, and re-iterable -- that is, any part of a collaborative analytic can be re-organized and re-structured around any other part. If you search in TAF on “ozone,” for example, you currently (July 2018) get 302 results that can then be filtered out according to type; most of those (179) are bibliographic references imported in from our collaborative Zotero group library, 65 are annotations, 27 are PDFs that have been uploaded directly to the site, and 1 is a question that is part of a collaborative analytic. Clicking one of the 65 annotations takes you to a particular response from a particular researcher to a particular artefact pertaining somehow to ozone (perhaps its effect on lungs); clicking from there on the collaborative question to which that annotation was written in response opens a page listing all the other responses to that analytic concerning ozone, by all other researchers, for all other artefacts (perhaps concerning its status in EPA regulatory policy, or

the environmental group that developed an app to provide near-real-time ozone data to Houstonians, or new satellite technologies for monitoring it). It's through such infrastructural design and developments that ethnographic collaboration on asthma (expansively understood), we hope, can be leveraged, multiplied, diversified -- re-organized and re-written experimentally, again and again.

2) The PECE essay is a digital-literary form also designed to instantiate and/or leverage such disjunctive, juxtapositional, collage-ish aesthetics and practices likened to surrealism by James Clifford (Clifford 1981). So taking any of those 302 search results for ozone, say, a TAF researcher can juxtapose on a single screen multiple images, documents, audio-video artefacts, or annotations authored by multiple collaborators, putting them into "lightly structured" relation (more or less loosely) to any chosen topic or subject area such as ozone. The PECE essay pursues those "kaleidoscopic logics" that are demanded by complex conditions such as asthma, and the other similarly disseminated, dissensus-riddled phenomena that demand collaborative work and involvement from ethnographers--but presented in ways that keeps the differences and "suturings" (Clifford's term) visible.

We thought this would be easy to code as digital infrastructure; it was not. We have invested extensively in software development for the PECE essay to make continual improvements: to allow multiple authors, for example, to allow annotations to be added to the essay, and to allow the overall essay itself to be annotated. Even after much time and money we are already dreaming of and pining for PECE Essay Version 1.2, and already know that we will then be wanting Version 1.3 which we can't as yet afford.

You can see an early example of Version 1.1 here: <http://theasthmafiles.org/content/mobilizing-asthma-research/essay>. A new collaboration which began after we first wrote this piece, STS Across Borders, makes extensive use of the PECE Essay to document and present the extensive research in science and technology studies and related activities occurring globally; UC Irvine anthropology graduate student Angela Okune, for example, curates and presents a wide range of materials pertaining to science, data, and science and technology studies in a number of African nations (Okune 2018): <http://stsinfrastructures.org/content/sts-africa-formation-1/essay>

By what means: Collaboration, to reiterate, requires infrastructure, media, a platform. When The Asthma Files project began fifteen years ago, it manifested publicly as a Powerpoint file and slideshow, but only on the occasion of a public talk by one or both of us. ("Public" in the sense of a small academic audience in a university classroom or auditorium.) As the project and number of collaborators grew that format quickly became unmanageable, and insufficient for documenting data provenance, for preserving and sharing data, and for similar "data life cycle" issues to which we were becoming more attuned through our involvement in the Research Data Alliance (see below). This led to the conceptualization, design, and development of PECE, with one of its aims being a more expansive "public" for the collaboration and collaborative research to address.

The Asthma Files research itself was for the most part squeezed into and out of our "spare time" for many years. Some of the undergraduate students who contributed research were paid as work-study students, but most did it for academic credit. The first time that dedicated funding and time for TAF-specific research was secured was for our "Six Cities" project, described next.

6+ Cities

Who: The core TAF collaboration (Kim and Mike; then Ali Kenner with her dissertation research) organized a research/teaching event on asthma and air pollution with Dan Price, a philosopher-turned-environmental studies scholar teaching at the University of Houston Houston. This built up the TAF research thread on "asthmatic spaces," which began to develop comparative studies of New York, Houston, and Tehran (through Tahereh Saheb's dissertation research). That thread of work has grown

into our current “Six+ Cities” project, also characterized by a research scope that has expanded from asthma (albeit “asthma” writ large) to become a comparative ethnographic study of environmental health governance in Philadelphia (where our lead collaborator remains Ali Kenner), Houston (Dan Price), Albany (Scott Kellogg), New York City (Sam Elrahman, Thomas Depree), Beijing (Rodolfo Hernandez), and Bangalore (Govind Gopakumar, Vinay Baidur). That collaborative project has continued to expanded even more into the open-ended “6+ Cities” project that now includes the additional Indian cities of New Delhi (Surajit Sarkar, Prerna Srigya), Hyderabad (Pankaj Sekhsaria, Aalok Khandekar), Chennai (Deepa Reddy), and Pune (John Mathew). This collaboration is anchored in an academic kinship network of our current and former students. Former graduate student (now an assistant professor of anthropology/sociology at IIT Hyderabad) Aalok Khandekar leads the 6+ Cities research in India, and our current graduate student Alli Morgan works on the group coordination and administration.

What: The collaborative research focus here shifted from TAF’s main concerns of asthma sciences and asthma care, to environmental health governance “styles” and how air pollution, in particular, is monitored, analyzed, and remediated (in the multiple senses of that term) in varied cultural/political contexts. Cities emerged as the chosen ethnographic object or level of analysis in part as an effect of a growing literature on data in “smart cities,” and its intersection with an increasing focus on cities in scholarship on governance. Our initial choice of six cities was intended to echo a landmark scientific collaboration of the 1980s that came to be called “The Six Cities Study,” one of many collaborative research projects in the sciences that we admire, and that have fueled our own collaborative drives. Most closely identified with the Harvard School of Public Health, the original Six Cities Study researchers (working in a different set of cities than ours) were the first to analyze large and diverse data sets to demonstrate conclusively that air pollution seriously harmed public health (Dockery et al. 1993).

When: As with the other collaborations described above, we’ve tried to plan regular weekly two-hour meetings with the local Rensselaer-based members, sometimes joined by others via skype. Beginning in 2017 we have tried to hold skype conference calls about every 2 or three weeks with most of the researchers in the India cities that were added. We have ourselves done a few two-week concentrated fieldwork stints of interviewing and research with our collaborator in Beijing Rodolfo Hernandez (Kim and Mike, November 2016) and some of our collaborators in India (Kim, March 2017), and similar briefer research stints with our U.S.-based collaborators. So this research, too, is regular and distributed over time, throughout the collaboration, coordinated largely through the TAF instance of the PECE platform.

Where: See the foregoing.

Why: Cities were chosen less by logic or through some research algorithm, and more by a kind of collaborative opportunism: where did we know someone, or know someone who knew someone, that had appropriate ethnographic interests or expertise? This in part explains the initial omission of seemingly obvious cities such as Los Angeles (now that we have moved to the University of California at Irvine, we are adding in Los Angeles and its neighboring areas). We are trying to say something about cities as a “middle” level or unit of governance, where air is understood through actions “styled” by multiple actors distributed across “lower” (city inhabitants, neighborhood groups and NGOs) and “higher” (regional, national, international bodies and regulations) levels.

In what way: Much of the research effort has been fluid and open-ended, a kind of operationalization of the “+” in the project’s title. As the TAF collaboration has grown into the 6+ Cities collaboration, though, there has been a need for more “light structure” to organize the research and its administration.

Especially in the case of the Indian cities, our collaborators have become in effect paid consultants, for whom we have tried to specify expectations and deliverables (e.g. 30 interviews, 100 entries added to a collective Zotero library, 20 article annotations, etc.). We've budgeted for 300 hours of work from each consultant, but as we state in our agreement with consultants: "We realize that research is never quite so straightforward...and will require more than 300 hours. The RPI group will help make up the difference. We also hope to create synergies that will help us accomplish more collaboratively than we could have individually. All project participants are free and encouraged to use material collected for the project in their own projects."³

By what means: Our initial six cities research was funded through a 2-year NSF research grant; a small percentage of this went to PECE platform development. The research in Indian cities that is part of the 6+ Cities project was funded through a grant from the Azim Premji Foundation in India.

Research Data Alliance's Digital Practices in History and Ethnography Interest Group (DPHE-IG RDA)

Who: Kim Fortun, Mike Fortun, and Jason Jackson (Indiana University, Mathers Museum of World Cultures, and openfolklore.org) are since 2013 co-chairs and co-founders of this "interest group" that is one of many and indeed a growing number of interest and working groups (e.g. a Metadata IG, a Wheat Data Interoperability IG, a Linguistics Data IG, a Structural Biology IG) that make up the Research Data Alliance.⁴ There are about 120 listed members of the DPHE-IG; at most about one-quarter of those are researchers that we know personally, who joined RDA through our interest group, the other 3/4ths are drawn from the broader RDA membership who have also subscribed to the DPHE-IG. Our graduate student and PECE Platform Architect Lindsay Poirier is now an "RDA Fellow," a recognition previously held by our former graduate student Brandon Costelloe-Kuehn and now also held by one of our newest collaborators, Vivian Wong (UCLA). RDA Fellows work with an IG and are remunerated with travel funds to the semi-annual Plenary Conferences held in cities like San Diego, Dublin, Amsterdam, Tokyo, and Barcelona. Our working sessions at RDA Plenary Conferences are usually attended by around 20-30 people; our "call-in" webinars in which researchers present their projects or discuss data sharing issues (these were organized most regularly in 2015-16 and more sporadically since) usually have 5-10 participants.

What: From the DPHE-IG description on the RDA website:

RDA's Digital Practices in History and Ethnography Interest Group (DPHP-IG) works to advance data standards, practices and infrastructure for historical and ethnographic research, contributing to broader efforts in the digital humanities and social sciences.

Goals

- Advance development of digital infrastructure for historical and ethnographic research through engagement with concrete scholarly practice and projects (such as Open Folklore, the Nunaliit Atlas Framework, the Platform for Experimental and Collaborative Ethnography and Indiana University's Mathers Museum of World Cultures).
- Advance conceptualization of the special characteristics and digital potential of humanities and qualitative social science data, including conceptualization of ethnographic and historical research data as "big data."

³ See <http://theasthmafiles.org/content/taf-6-cities-work-plan-project-consultants>)

⁴ <https://www.rd-alliance.org/groups/digital-practices-history-and-ethnography-ig.html>

- Advance capacity to share, integrate, visualize and act with different kinds of data and analyses, including qualitative data and the kinds of analyses produced through historical and ethnographic research.

Planned Outcomes & Benefits

- Build a global network of people involved in the development of data infrastructure for historical and ethnographic research, providing opportunities to share digital tools and project development experience. Monthly, call-in “project shares” since summer 2013 contribute to this.
- Link people involved in development of data infrastructure for historical and ethnographic research to data scientists and technologists, and to people in other research domains involved in data infrastructure development (leveraging the connections provided by RDA).
- Characterize and recommend best-practice meta-data standards for researcher-created primary data (field notes, recorded interviews, etc.) in history and ethnography. This will be the focus of the first Working Group spun out of this Interest Group.
- Characterize and recommend user agreements, citation practices, digital exhibition protocols, and other mechanisms that will facilitate sharing and public availability of historical and ethnographic data (recognizing the need to customize access according to data type and context).
- Develop an ethnographic project to document and analyze data practices and culture in different research communities, especially as represented in the RDA. The comparative knowledge created by the project can undergird deep research collaboration across diverse fields.

When: RDA formed in 2013, and our interest group was one of the early IGs formed that year--the same year that the U.S. formalized its Federal Open Data Policy requiring government data be made available in open, machine-readable formats. It was developments like these in both government and scientific data landscapes that reinforced our sense that infrastructure was vital to support; we found ourselves both studying the “open data” wave as researchers, and simultaneously riding it as open data practitioners

This more distributed, looser collaboration runs mostly in the background for us, but makes productive sporadic foregroundings in the forms of the virtual “project share” GoToMeetings (for a while about once every few weeks, now less frequent), and when one or two of the co-chairs or our RDA Fellows attend an RDA Plenary Conference for face-to-face working sessions.

Where: The DPHE-IG originated at RPI when we worked there; it is not completely coincidental that key RDA leadership figures (in computer and web sciences) are also at RPI. RDA styles itself as “global” (5,400 members in 123 countries), although there are three main divisions: RDA-US, RDA-EU, and RDA-AU. The biannual plenary conferences have been held almost entirely in major cities in the U.S. and EU: Dublin, Washington DC, Amsterdam, San Diego, Paris, Berlin. Asia and Africa are acknowledged as places where outreach and greater inclusion need to occur; Tokyo was the site of one of the plenary meetings in 2016, and Gabarone, Botswana is the site of a 2018 plenary meeting.

Although most of our collaborative work with RDA occurs virtually, as is true of RDA generally, our group meetings at the Plenary Conferences have been important not only for keeping abreast of new data developments but also for making new contacts and beginning new collaborations; UC San Diego computer scientist Ilya Zaslavsky, for example, attended our IG session at the Berlin conference in March 2018, leading to a conversation that has grown into a new collaboration that will try to make

the more qualitative methods and tools of PECE interoperable with the more statistical quantitative methods and tools he has designed called SuAVE (Survey Analysis via Visual Exploration, <http://suave.sdsc.edu/>)

Why: The DPHE-IG adopts the presumptive promise of RDA itself: that data sharing is inherently good, and furthering the freer and more open sharing of more (and better) data is part of a researcher's ethico-political responsibility. A primary way through which that responsibility is manifested is through the construction of and care for digital infrastructure (including cultural protocols) for the creation, archiving, maintenance, and sharing of data. Of course, matters of privacy and the ethics and regulation of data flows and use are subjects of RDA interest and work, but openness, sharing, and borderlessness are the dominant tropes:

The Research Data Alliance (RDA) was launched as a community-driven organization in 2013 by the European Commission, the United States National Science Foundation and National Institute of Standards and Technology, and the Australian Government's Department of Innovation with the goal of building the social and technical infrastructure to enable open sharing of data.

With over 7000 members from 137 countries (June 2018), RDA provides a neutral space where its members can come together through focused global Working and Interest Groups to develop and adopt infrastructure that promotes data-sharing and data-driven research, and accelerate the growth of a cohesive data community that integrates contributors across domain, research, national, geographical and generational boundaries. (<https://www.rd-alliance.org/about-rda>)

In what way: RDA supplied crucial funding for the development of PECE data infrastructure; we were awarded a grant to adopt the "outcomes" of RDA's Practical Policies Working Group by coding them into the data model, data handling, and permissions systems of PECE. RDA has also been a constant source of comparative insights and ideas from "data wranglers" in numerous fields, from neutron physics and genomics to library and information sciences.

By what means: The DPHE-IG collaboration works mostly virtually (GoTo Meeting, skype conferencing, Google docs). When we first started, we hosted bi-weekly project shares where we invited researchers with noteworthy experience in data sharing efforts in any domain (e.g. [Scalar](#) and [Project Bamboo](#) in the digital humanities, the [National Snow and Ice Data Center](#) and [DataONE](#) in the sciences. These have become less frequent, and our work has in large part shifted toward a spin-off working group (the Empirical Humanities Metadata WG) that will survey, summarize, and distill preferred metadata models and practices in what we call the "empirical humanities:" ethnographic and historical research that has as a primary goal the generation of new data, that will be further enhanced through "collaborative hermeneutics" and shared widely for continued re-interpretation and re-datafication. The WG "Case Statement" (which had to go through multiple rounds of editing and review before the group was approved) lays out the rationale:

Given the cultural and social complexity (as well as technical, ecological and economic complexity) of many global problems today, collaborative empirical humanities research has renewed urgency. For decades, research in these fields has been an almost entirely individual-centric enterprise. Field notes, found documents, found or researcher-created photographs or recordings and other data used in cultural analysis are very rarely shared, except when reduced or rendered into some form of publication or museum display.

One of the primary barriers to sharing data within the empirical humanities is a lack of agreed-

upon protocols for metadata standards for user-created primary research data. While there has been a great deal of work in the cultural heritage arena, especially within museums and libraries, and the dilemmas of qualitative data re-use are well documented (see Holstein and Gubrium 2004), the issues associated with preparing data for later use by third parties are yet to be thoroughly conceptualized...Many researchers find themselves caught in the confusing space between the dizzying proliferation of standards and a one-size-fits-all approach that can miss out on the diversity of data practices within disciplines. Working closely with existing metadata-focused RDA groups...we will produce a simple list of recommended metadata fields for a delimited set of artifact types, analytics and use cases. Once endorsed by the RDA, and taken up by early adopters, these best practices will be a go-to resource for researchers that may then choose to modify (add or subtract) the fields we suggest for their own purposes. Development and uptake of shared metadata practices and tools will make user-created research data more findable and usable within these research traditions. The work of this WG could also contribute to the development of mechanisms providing greater credit and incentives for sharing data.⁵

**Conclusion: The Depositivist (Feverishly Archival, Experimental, Infrastructural)
Style of Collaboration**

In our 6+ Cities Project, we have come to summarize a key goal of our collaborative efforts as explicating the “environmental health governance style” of each city: how a city’s diverse actors (agencies, experts, citizens) comes to know and tries to improve the composition of the air over time in that region. How could we characterize the mixes of democratic, bureaucratic, and technocratic patterns and movements peculiar to each place, with its variable nestings in local, regional, national, and global institutions, regulatory regimes, and histories? How do citizens get involved (or not), how do sciences get understood and referenced (or not), how are technical and socioeconomic resources invested (or not)?

There were several reasons we chose “style” as a kind of governing concept for our research: for one, we found long-standing efforts by historians, sociologists and philosophers of science (and scientists themselves) to characterize “styles” of thinking, reasoning, or “doing science” to be interesting and worth extending. Ludwik Fleck’s concept of the “thought-style” is perhaps most familiar, as is Ian Hacking’s extensive work characterizing different styles of scientific reason (summarized in Hacking 2012). Like “culture,” “style” connoted for us something subtle and sub-surface yet also substantial, elusive yet omnipresent, and enduring yet inviting of change and play. The experimental style, speculative style, deductive style, or statistical style of doing science, to name but a few of the styles that have been traced, have been fruitful designations to differentiate and qualify a “science” otherwise prone to unification, idealization, or reification, while avoiding complete avoidance of the “science” category. Strategic essentialism by stylization, if you will, or writing science under (styled) erasure.

And that was another reason for adopting “style:” as a sign of something that, no matter how subtle or elusive or shape-shifting, is always somehow written (with a *stylus*). To be styled -- and only science (at least a certain version of it) dreams of itself, or its future self, as unstyled -- is to be impressed in or into a material or medium: a piece of paper, a body, a cityscape, a distributed digital network. Each of our collaborations was, in this sense, a project of writing/styling -- not just writing about, and not just involving writing as one of the activities that collaborators did, but itself written/styled.

⁵ <https://rd-alliance.org/group/empirical-humanities-metadata-working-group/case-statement/empirical-humanities-metadata>

Collaboration too, like science, might be fruitfully approached and analyzed through the concept of style, then, rather than along its more conventionally codified axes (cooperative/competitive, social/solitary, gift/commodity, etc.). How would we characterize the style of our collaboration? We start by naming it the depositivist style, and go on to elaborate its feverishly archival, experimental, and infrastructural qualities that together compose our style of collaboration.

Our depositivist style of collaboration is marked by the trace of a positivist style of science, but one with its ground mined under by the play and work of deconstruction. To name only one sign of this aspect of the depositivist style: our comfort with and even embrace of the term “data,” which raises more than a few hackles among more than a few anthropologists. (Our essay’s title is a wry reference to these kinds of responses, that find our affinities to the sciences and their ideals somewhat funny, odd, or otherwise out of the anthropological mainstream.) The depositivist style also embraces the experimentalist style, as described earlier in this essay, that characterizes even the most positivist of sciences. Depositivism thus aligns our collaborative work in anthropology in a broader conception and history of the human (i.e. styled) sciences.

As is true of so many contemporary sciences, our depositivist style of collaboration is also one which privileges the deposition or archiving of data as much as and in many cases more its use. Depositivism is a style of deferral, then, another trace of its broader deconstructivist legacy. It is a sedimentary style of collaborative anthropology in the same manner making up related scientific styles of work and thought:

The metaphor of sedimentation comes from Husserl’s *Crisis* (1970), but is used in a way opposite to his. In caricature, Husserl thought it was the philosophical task of his time of European crisis to clear away the sediment. That was a specific response to the 1930s. We should, he thought, try to recover the original experiences underlying fundamental events, such as the discovery of mathematical proof (the *Ursprung* of geometry) and Galileo’s mathematization of nature. He wanted to remove the sediment, to reach, perhaps, a pre-Galilean state. In our contrary perspective, the sediment, hardened over a long time by great pressures into rock, is a collection of achievements founded on human ingenuity, innate propensities, and interaction with everything. Like any sedimentary deposit, it may undergo radical change in the future, but it cannot be undone. (Hacking 2012: 600)

The depositivist style of archiving is a feverish one, however, a heated and ill-at-ease one that differs from the calmer assurances of positivist archives. Jacques Derrida’s *Archive Fever*, a short book styled as a long essay that has impressed itself upon our collaborations in many ways, elucidates this aspect of our style. Like nearly all of Derrida’s writings, this essay tracks back and forth through a tight series of tangles concerning psychoanalysis, its history, history and memory more broadly, science and techno-science, and the question of whether psychoanalysis could or should be named--*tagged*, we might say today, in our new tele-technological landscape--a “Jewish science.” Derrida too refuses the dream of an un-styled science, sciences un-impressed by the conditions of their production, affirming psychoanalysis’ status as a science (contra Karl Popper), but a styled, “Jewish”(-ish) one. In process he raises the question of archives, and the feverish, troubled un-easiness (*mal d’archive*) of their authority and status:

[T]he question of the archive is not, we repeat, a question of the past. It is not a question of a concept dealing with the past that might *already* be at our disposal or not at our disposal, an archivable concept of the archive. It is a question of the future, the question of the future itself, the question of a response, of a promise and of a responsibility for tomorrow. The archive: if we want to know what that will have meant, we will only know in times to come. Perhaps. Not tomorrow but in times to come, later on or perhaps never. A spectral messianicity is at work in

the concept of the archive and ties it, like religion, like history, like science itself, to a very singular experience of the promise. (Derrida 1998: 36)

For us, then, the depositivist style is only partly about what a collaboration has actually achieved or archived--although we are not devoid of or disinterested in such deliverables. The depositivist style is even more about what the experimental, archival collaboration is becoming, or will have become, and what infrastructures can keep that archive becoming, keep it troubled and feverishly excited and excitable, and experimentally open to un-planned futures. Depositivist collaboration is collaboration for the sake of building (infrastructure for) further, more robust collaboration. The depositivist style is a collaborative style directed only in part by a defined collective end product, tangible result, culminating exhibition, or project; it is even more attentive to organizing the ongoing work of a collective towards continually reiterating itself, extending itself into new collaborations.

A depositivist (archival, experimental, infrastructural) style is a promising style, requiring trust from its collaborators, and commitments of care -- always putting the labor back into collaboration. It is patient and feverish, oscillating between satisfied and dissatisfied with the interminable work of building collaborative infrastructure for future collaborations that might and might not be there. A depositivist style is an uneasy and responsive style, a style of stay-with-us and stay-with-it, a style of impatient deferral that keeps on keeping promises...

References

- Brown, Susan, Patricia Clements, and Isobel Grundy. "Sorting Things in: Feminist Knowledge Representation and Changing Modes of Scholarly Production." *Women's Studies International Forum*, Feminisms and Print Culture, 1830-1930s, in the Digital Age, 29, no. 3 (May 2006): 317–25. doi:10.1016/j.wsif.2006.04.010
- Clifford, James. 1981. "On Ethnographic Surrealism." *Comparative Studies in Society and History* 23(4): 539–564. www.jstor.org/stable/178393.
- Derrida, Jacques. *Archive Fever: A Freudian Impression*. University of Chicago Press, 1998.
- Dockery, Douglas W., C. Arden Pope, Xiping Xu, John D. Spengler, James H. Ware, Martha E. Fay, Benjamin G. Ferris, and Frank E. Speizer. "An Association between Air Pollution and Mortality in Six U.S. Cities." *New England Journal of Medicine* 329, no. 24 (December 9, 1993): 1753–59. doi:10.1056/NEJM199312093292401.
- Fortun, Kim, and Mike Fortun. "An Infrastructural Moment in the Human Sciences." *Cultural Anthropology* 30, no. 3 (August 10, 2015): 359–67. doi:10.14506/ca30.3.01.
- Fortun, Mike, Kim Fortun, and George Marcus. 2016. "Computers in/and Anthropology: The Poetics and Politics of Digitization," in Larissa Hjorth, Heather Horst, Anne Galloway, and Genevieve Bell (eds.), *Routledge Companion to Digital Ethnography*. New York: Routledge, pp. 11-20.
- Gitelman, Lisa, ed. *Raw Data Is an Oxymoron*. Cambridge, MA: MIT Press, 2013. <https://mitpress.mit.edu/books/raw-data-oxymoron>.
- Gold, Matthew K. *Debates in the Digital Humanities*. U of Minnesota Press, 2012.

Hacking, Ian. 2012. 'Language, Truth and Reason' 30 Years Later. *Studies in History and Philosophy of Science Part A* 43(4). Part Special Issue: Styles of Thinking: 599–609.

Kenner, Alison. 2018. *Breathtaking: Asthma Care in a Time of Climate Change*. Minneapolis: University of Minnesota Press.

Klein, Lauren F. "The Image of Absence: Archival Silence, Data Visualization, and James Hemings." *American Literature* 85, no. 4 (December 1, 2013): 661–88. doi:10.1215/00029831-2367310.

Marcus, George E., and Michael M.J. Fischer. 1986. *Anthropology as Cultural Critique: An Experimental Moment in the Human Sciences*. University of Chicago Press.

Turin, Mark. "Born Archival: The Ebb and Flow of Digital Documents from the Field." *History and Anthropology* 22, no. 4 (December 1, 2011): 445–60. doi:10.1080/02757206.2011.626776.

What's so funny 'bout show more As I walk through This wicked world Searchin' for light in the darkness of insanity. I ask myself Is all hope lost? Is there only pain and hatred, and misery? And each time I feel like this inside, There's one thing I wanna know: What's so funny 'bout peace love & understanding? What's so funny 'bout peace love & understanding? And as I walked on Through troubled times My spirit gets so downhearted sometimes So where are the strong And who are the trusted? And where is the harmony? Sweet harmony. 'Cause eac