Organizing for ontological change: The kernel of an AIDS research infrastructure

David Ribes
Communication, Culture & Technology, Georgetown University, Washington, DC, USA

Jessica Beth Polk
Department of Science & Technology Studies, Cornell University, Ithaca, NY, USA

Abstract
Is it possible to prepare and plan for emergent and changing objects of research? Members of the Multicenter AIDS Cohort Study have been investigating AIDS for over 30 years, and in that time, the disease has been repeatedly transformed. Over the years and across many changes, members have continued to study HIV disease while in the process regenerating an adaptable research organization. The key to sustaining this technoscientific flexibility has been what we call the kernel of a research infrastructure: ongoing efforts to maintain the availability of resources and services that may be brought to bear in the investigation of new objects. In the case of the Multicenter AIDS Cohort Study, these resources are as follows: specimens and data, calibrated instruments, heterogeneous experts, and participating cohorts of gay and bisexual men. We track three ontological transformations, examining how members prepared for and responded to changes: the discovery of a novel retroviral agent (HIV), the ability to test for that agent, and the transition of the disease from fatal to chronic through pharmaceutical intervention. Respectively, we call the work, ‘technologies’, and techniques of adapting to these changes, ‘repurposing’, ‘elaborating’, and ‘extending the kernel’.

Keywords
AIDS, data, historical ontology, HIV, infrastructure, instruments, kernel, objects of research, specimens, subjects

Corresponding author:
David Ribes, Communication, Culture & Technology (CCT), Georgetown University, 3520 Prospect St. NW, Suite 311, Washington, DC 20057, USA.
Email: david.ribes@georgetown.edu
**Introduction**

We have come to understand ontological change as that which emerges from scientific investigation, often in a surprising or unexpected manner. In the lab, it is an anomalous result or outlier, first to be shelved and then insisting on explanation. In observation, emergence is a cognitive awakening, a new habit of seeing, or a shift in perception. Or, in discourse, it comes about through novel categories, processes of naming, and new representational techniques such as imaging or quantification. In such formulations, new entities are to be incorporated into scientific investigation, found retrospectively as always already present in the past, and prospectively managed as a component of all future research.

But ontological change is something that can be planned for and built into the design of a study or research infrastructure. While the technical details and content of ontological change are emergent, in broad strokes they can be anticipated and fertile circumstances can be arranged in advance. In this article, we identify actors’ methods for preparing, planning, and responding to ontological change. For a long-term study – built to continue over the course of years or decades – ontological change was expected. In the case we examine, change was desired, sought out, and enacted.

The Multicenter AIDS Cohort Study (MACS) was founded in 1983, tasked with identifying the cause and patterns of transmission of a very poorly understood disease that appeared to attack narrow demographics with deadly outcomes: AIDS. Today, the MACS continues as the longest running study of HIV, the retroviral agent that, unchecked, leads to AIDS and likely death. However, in 1983, HIV had not yet been discovered as the cause of AIDS. While an infectious agent was suspected, other explanations were under investigation within medical circles. To investigate the cause and modes of transmission, the MACS was founded in four American cities, and a cohort of nearly five thousand gay and bisexual men was recruited for a prospective epidemiological investigation. Today, over thirty years later, and across many changes to AIDS science, MACS members continue to investigate HIV disease while regenerating an adaptable research organization.

Our central interest in this article is to examine MACS’ ability to prepare, respond, and adapt to changes in its research objects while remaining ‘the same study’ – a quality we call technoscientific flexibility. The key to this flexibility is the preparatory design, and, thereafter, adaptation of what we call the *kernel of a research infrastructure*. Our inspiration for this concept stems from computing. In software engineering, a kernel is the main component of a computer’s operating system (OS), facilitating access to resources and services. The kernel is the infrastructure of practical computation, creating an abstraction layer that, for programmers, greatly facilitates access to system resources.

Similarly, the kernel of a research infrastructure seeks to facilitate access to the resources and services needed to investigate scientific objects. In its thirty years, the MACS kernel has been designed and redesigned in preparation for expected changes and retrospectively reorganized to adapt to emergent transformations. We examine three moments from the history of the MACS, highlighting how members prepared and responded to ontological changes in their objects of study: the discovery of a novel retroviral agent in 1984 (eventually called HIV); the ability to test for that agent in 1985, shifting participating subjects from being ‘at risk’ to being HIV-positive or HIV-negative;
and the transition of the disease from fatal to chronic through pharmaceutical intervention in 1996. Paradoxically, even as changes to the kernel made it possible to study certain objects, investigating other objects became increasingly more challenging.

While we examine the adaptability of the kernel in the face of ontological change, persistence is the kernel’s most important quality – such as the biannual trip participating men take to clinics in four US cities, contributing their time and bodies that are then translated into long-term data and specimen archives. Together, we call these reflexive activities, persistent routines, and material rearrangements, organizing for ontological change.

The kernel of research infrastructure and ontological change

Environments that sustain controlled chaos are the heartland of innovation. (Timothy Lenoir)

We adopt a circumscribed definition of ontological change. In this article, ontology refers specifically to the objects of investigation, and change to how they emerge, fade, or recur as objects of investigation. The kernel is the constellation of concepts we employ to explain technoscientific flexibility: the capacity of research infrastructure to support investigations of partially or wholly unexpected objects. In addition to ontological changes, we also examine key moments of change to the kernel that sought to enable the investigation of new objects. Although we do track such changes, the kernel’s primary characteristic is its continuity and, counter-intuitively, it is the ongoing, sustained availability of the kernel’s resources and services that explains its ability to adapt to ontological change.

Our approach is in the tradition of historical ontology (Hacking, 2004: 1–22), examining the ways that objects of research change over time, including how these changes occur as a matter of scientific intervention. Our understanding of ‘objects of research’ is broad, tracking what investigators themselves identify as their objects and the activities and materials used to craft them as such. Following Susan Leigh Star (2010) in the pragmatist sociological tradition, the generation of objects ‘derives from action, not from a sense of prefabricated stuff or “thing”-ness’ (p. 603). Objects may be concrete, such as Kaposi’s sarcoma tumors, or more abstract, such as the natural history of AIDS itself; they may be of biological interest, such as CD4 counts, or of social scientific focus, such as mood, coping, and stress; they may be entities such as glycoproteins gp120 and gp41 that line the envelope of the HIV retrovirus, or processes and relations such as reverse transcription and aging with HIV. Our own central interest is in understanding how members organize their material environment and practical expertise to manage partially or wholly unforeseen transformations to their objects of research.

The kernel of a research infrastructure is made up of the resources and services that members work to keep available in preparing, managing, and responding to changing research objects, all the while continuing to support ongoing investigations. The kernel is partially designed with flexibility in mind. It is assembled with a view to the future, serving to continue the study of extant objects while keeping tabs on certain resources that
may come to be useful for investigations of new and unforeseen research objects. However, the kernel is not only designed, but it is also shaped by accretion and shedding over time, often in the form of ongoing maintenance, repair, and upgrade to its components.

If misconstrued, the term ‘kernel’ can be taken the wrong way: a seed that already contains all the information needed to grow into corn. Instead, our inspiration stems from computing, where a kernel is the main component of an OS. On its own it does little, but for an application programmer the kernel greatly facilitates access to system resources such as processing, storage, and peripheral devices. A programmer need only have a simplified understanding of how to access resources (e.g. memory) because the task of managing such operations is delegated to the kernel. Similarly, in conducting a scientific study, a particular investigator need not be fully aware of the full range of activities and operations that have made resources available. For example, in using a blood specimen collected by the MACS, a researcher may know very little about the carefully calibrated collection of that specimen – orchestrated to ensure comparability with specimens collected over 30 years and across four geographic sites.

Within computer engineering, an OS kernel is most often visually represented in a simplified ‘layer cake diagram’ or ‘stack’. This is a useful representational convention for communicating a complex structure, and we adopt it for heuristic purposes (see Figure 1 for the MACS kernel as a stack). However, encountered as a sociotechnical assemblage (Kelty, 2008), no single human actor can understand a kernel’s full intricacies. An OS kernel may be millions of lines of code worked on by hundreds of programmers, sedimented over decades. The kernel is not a static thing; it is versioned. There are actors devoted to its maintenance, retrospectively repairing its operations in response to feedback, and prospectively upgrading its functions in anticipation of novel computational capacities. This prospective capacity is called ‘extensibility’: design-oriented to support imagined future needs. Such upgrades operate in tension with goals for ‘backwards compatibility’; that is, extra efforts must be invested into ensuring the continued functioning of legacy devices, data, or software.

The kernel of a research infrastructure is not defined by always already containing all the resources needed to examine a new object; it only refers to what members expect may be of importance in the study of extant and expected future research objects. In the case of the MACS, there are four primary resources: (a) archives of specimens and data, (b) sites for the collection of those materials (the human participants), (c) calibrated instruments, and (d) heterogeneous experts. Each of these four resources and services are transformed and renewed to prepare and respond to changing objects of investigation. Other research infrastructures reveal altogether different kernel architectures (Ribes, 2014): which resources and services are assembled and sustained are profoundly entangled with present, past, or expected objects of investigation. Thus, we expect that the methodology described here can be extended to other research infrastructures by attending to the configuration of resources and services sustained within those organizations.

Preparing and responding to ontological change: the role of infrastructure

Generally, in Science and Technology Studies (STS), use of the term ‘emergent’ is a contestation of any approach that takes objects as readymade in nature. ‘Emergent’
means that the object of study is surprising and recalcitrant to understanding and prediction, that the object is crafted in practice at the intersection of instruments, technique, and over time (Rheinberger, 2010). However, in this article, we are less interested in how or even why new objects emerge, but rather, with the ways in which the sciences prepare for new objects – a focus which has received far less attention.

Such a focus necessarily broadens our attention from the hands-on activities with scientific materials to include, say, the hands-on activities of laboratory record-keeping in notebooks. For instance, recounting how a serendipitous encounter of a bacterial culture with heat led to the first attenuated cholera microbes, Bruno Latour (1983) quipped that ‘chance favors only well-prepared laboratories’ (p. 148). He was referring to the ordered lab notebooks that had made it possible to track-back to what had occurred to attenuate those bacilli. Our analysis continues in this trajectory, investigating the material and organizational infrastructures of science by turning our attention to ongoing activities of data collection, curation, and reuse (Hine, 2006; Leonelli, 2012), to technologies that support coordination and collaboration (Jirotka et al., 2013), and to a socio-technical organization that facilitates the addition of new specializations, instruments, and research questions.

This is a form of ‘infrastructural inversion’ (Bowker and Star, 1999): rather than first examining research or scientific findings themselves, inspecting the kernel begins with
those resources and services that make research possible. This may include seemingly mundane work like instrument calibration or specimen procurement (Clarke and Fujimura, 1992), quality control (Petty and Heimer, 2011), and the creation of documentary forms to render activities accountable (Jensen and Winthereik, 2013) or to facilitate data sharing across teams of researchers (Edwards et al., 2011; Karasti et al., 2010). Such tasks are often backgrounded in accounts of science because they are developed, implemented, and sustained by technical workers and administrative staff (Millerand et al., 2013; Monteiro et al., 2013), and delegated to computational systems (Ribes et al., 2013). However, it is these activities that sustain availability by prospectively and retrospectively responding to change.

It is from such studies of infrastructure that the kernel of a research infrastructure draws its intellectual lineage. Concepts such as ‘standardized packages’ (Fujimura, 1992) or ‘platforms’ (Keating and Cambrosio, 2003) point to the entanglement of data, materials, methods, and instruments, as broad technoscientific assemblages form around key scientific domains and approaches (e.g. molecular biology or immunophenotyping). The concept of the kernel is simultaneously narrower than a standardized package or platform because it focuses on ‘an infrastructure’ rather than a field or domain; but it is also broader in scope, drawing the analysts’ attention not to a single instrument or approach, but to all the resources and services made available by that infrastructure across its history. An analysis of a kernel pushes past an exclusive focus on data that increasingly appears to be shared by both infrastructure practitioners and its scholars, and broadens the analysis to include material specimens (Landecker, 2007), where those data and specimens come from, and the instrument architectures that ensure their comparability (Waterton et al., 2013).

Lastly, a kernel is not tied to supporting a single style of scientific practice (Fujimura and Chou, 1994) but can be creatively repurposed to serve many epistemic cultures (Knorr-Cetina, 1999) and enact multiple ontologies (Mol, 2002). For example, MACS’ resources have been assembled to conduct epidemiological investigations of populations, laboratory studies of microorganisms, or social scientific investigations of motivations and affective states. In studying a kernel, the analyst must track its application with an agility matching the disciplinary range of its uses.

Changes to the kernel

In general, the kernel is the persistent set of resources sustained over time. In many circumstances, new objects of investigation can be assembled without demanding change to the kernel: what we call repurposing. Our primary example of repurposing is the discovery of HIV. Drawing from its existing kernel, the MACS was able to continue supporting investigations largely unchanged despite this key ontological change in the landscape of AIDS science.

However, in addition to sustaining availability of resources for reuse, the kernel also has the ability to facilitate change to itself – it is static in the sense that there are actors dedicated to maintaining its continuity, but also dynamic in that it is reflexively operated on by organizational members. We examine two changes to the kernel that we call elaboration and extension. Elaboration refers to the introduction of new instruments...
or analytical categories that bring additional depth to objects of research. We use the development of the HIV test to examine elaboration. In this ontological transformation, MACS participants shifted from ‘at risk’ of having AIDS, to being either HIV-positive or HIV-negative – a pivot point for new medical interventions, for care of the self, and, eventually, for new identities (Hacking, 2000).

Extension refers to the addition of resources and services to the kernel that enable the analysis of new objects. The transformation of HIV disease into a chronic illness serves to explore how extension works in practice. As HIV became chronic rather than fatal, the primary mission of the MACS shifted, from a natural history of AIDS to the treated histories of HIV. With this change came a plethora of new objects, never before investigated in biomedicine, such as aging with HIV.

In both cases of change, we will see that the novel resources and services enabling these emergent objects of investigation cannot simply be ‘dropped into’ the kernel. New resources go through a period of calibration, standardization, and testing that seeks to ensure backward compatibility with other resources. Changes to the kernel must be coupled with efforts to ensure continuity and interoperability, revealing a sought-after modular quality to the kernel. We discuss how a new instrument, the HIV test, was added in ways that sought to ensure similar implementation across the MACS’ four geographical sites and into its future.

The three ontological changes that we track in this article closely reflect the categories that members and participants themselves use to parse the history of the MACS; this said, we have also theoretically sampled (Glaser, 1978) these moments so as to target conceptual development of repurposing, elaborating, and extending of the kernel. We investigated the MACS through interviews and ethnographic participation in contemporary studies. Our more granular view of ontological changes comes from inspecting the archive of MACS publications, formal reports, and requests for proposals, tracing year by year as new objects of research were introduced or failed to recur in the study. We have also inspected data archives, which reveal – sometimes annually – changes in categories, methods, instruments, and commensurations across those changes (Bowker and Star, 1999: esp.Ch. 2). Additionally, the MACS website is an archive filled with seemingly mundane materials such as stock text for publication, slide decks, and data and material transmission forms. Thus, our research is bounded by the activities of MACS investigators over time, rather than, for example, being coextensive with the many fields that make up research on HIV/AIDS. In its thirty years, the MACS has undergone many other forms of change – for instance, in how it is funded, in the composition of its leadership team, or in its data sharing policies (Ribes and Polk, 2014) – but here we will focus only on changes to its objects of study and its kernel.

The bounded technoscientific flexibility of the kernel

Perhaps the greatest value of inspecting infrastructure through its kernel is that it enables a concrete analysis of what scientists can or cannot investigate by drawing on its specific resources and services. Returning to our driving concept, an OS kernel enables and constrains the actions of an application: If programmers do not follow its set of rules, resources will not be made available, the application has an error, the system may crash.
But within its constraints, the kernel facilitates the allocation of resources as programmers please, generating an infinite set of emergent applications.

Just as applications are enabled, but not cast in stone by an OS kernel, so too is the investigation of objects of research dependent on, but not determined by, the kernel of a research infrastructure. The kernel facilitates a bounded technoscientific flexibility, enabling the investigation of certain objects while restricting others. Ontology is emergent, but it is not de novo. Given a set of resources, certain objects will be readily available for analysis while others will be difficult or impossible to examine. Michelle Murphy (2006) has called this a ‘domain of imperceptibility’. In her study of Sick Building Syndrome, that potential object was repeatedly contested as it remained below a threshold of scientific and regulatory perceptibility. Such objects then inhabit what Law and Lien (2013) call a ‘penumbra of not quite realized entities’. As Lorraine Daston (2000) notes, ‘in contrast to quotidian objects, scientific objects are elusive and hard-won’ (p. 2). We see several such objects in the history of the MACS, objects that cannot be assembled from its kernel and consequently cannot be investigated.

The kernel can be considered a kind of archive, in the sense of the term employed by Geoffrey Bowker (2006). As with any archive, it comprises practices of inclusion and exclusion, the results of which propagate downstream over its history. The data and specimen repositories are literal archives – what is preserved is a matter of scientific deliberation, technical capacity, and pragmatic questions of information management and sample curation. The expert base and the subject cohorts are more figurative kinds of archives: throughout their history, MACS members have made decisions about what subjects should supply to archives and what scientific specializations can legitimately contribute their expertise in support of research. Paradoxically, each exclusion may place some objects within a domain of imperceptibility while simultaneously strengthening the ability to investigate others.

Building the MACS kernel: finding the cause

You know, we didn’t know what the cause of this was. (John P Phair, MD, PI Chicago and Chair of the MACS Executive Committee)³

Almost all historical accounts of AIDS trace the first medical announcement to the Centers for Disease Control’s 5 June, 1981 ‘Morbidity and Mortality Weekly Report’ (MMWR). That paper outlined many of the traits that thereafter informed efforts to characterize the disease and its cause. Of the five men identified in MMWR, all were young, ‘active homosexuals’, who had cytomegalovirus (CMV) and *Pneumocystis carinii* pneumonia (PCP). An immune disorder was quickly suspected because of these opportunistic infections, and because three of the men had markedly diminished T-lymphocyte populations. A second report in July added Kaposi’s sarcoma and several other diseases to the list (Centers for Disease Control, 1981). A proven track record of investigating these objects, using these instruments and methods, and having access to these populations proved critical in the National Institutes of Health (NIH) award decisions that became the MACS. As historians Fee and Fox note, most scientists and doctors at the time looked to epidemics of infectious diseases, and specifically, polio, yellow fever, or cholera, as a
model to make sense of AIDS. The standard public health response was ‘enhanced surveillance, mobilization of medical resources, and increased research’ (Fee and Fox, 1992: 3). Cohort studies were a key part of this strategy.

In this section, we examine the founding period of the MACS, with a focus on the assembly of its kernel. We treat each component of the kernel in turn, beginning with the scientific, medical, and institutional actors that came together to form the expert community of the MACS; we then turn to the choice and acquisition of instruments; the planning for a data and specimen repository; and, finally, the selection and recruitment of a study population. Notably, all of these actions occurred before the identification of any causal agent for AIDS, necessitating a ‘collect anything and everything you can think of’ approach in this budding investigation. While the MACS kernel is certainly broad, by necessity it also comprises decisions to target particular resources and services, simultaneously excluding others: in this sense, a kernel is built with a bounded flexibility in mind. We explore the implications of such decisions by briefly examining a controversy during MACS’ initial cohort recruitment.

The experts

The NIH – in particular, the National Institute of Allergy and Infectious Diseases and the National Cancer Institute – funded four sites that became the MACS in 1983. The initial research questions driving the projects were set out in the NIH request for proposals (RFP): the goal of the study was to investigate the cause and modes of transmission for AIDS; the design and methods were a prospective cohort study of gay and bisexual men in combination with laboratory studies. The day, 30 September 1983, marked the award date for the four selected Principal Investigators (PIs) in Chicago, Baltimore/Washington D.C., Pittsburgh, and Los Angeles. In total, NIH awarded five contracts. The fifth site never joined the MACS and became the San Francisco Men’s Health Study (SFMHS), which we discuss below. These PIs acted as the study’s first bases of expertise: medical doctors grounded in epidemiology, infectious disease, immunology, and one PhD in virology.

At its funding and as a result of how the RFP had been conceived and written, the studies were not multicenter in design; they were individual awards. None of the coordinating work usually included in multi-sited projects was outlined or budgeted. The decision to form a single study came from the PIs themselves: ‘The first thing we decided was we would form a multicenter study’. Thus, the first months of the study were logistical. Before beginning recruitment of subjects, MACS PIs and NIH program officers needed to determine the subject recruitment protocol, what specimens were to be collected and preserved, the questionnaires to be administered, how and where data would reside, and the methods for curating them:

A multicenter study is a logical way for [an] epidemiology study like [the MACS] to be done. You need the numbers [of subjects], and you want to … compare the data. You want to look at what’s different in Pittsburgh compared to LA or to Baltimore, or Chicago … And so to combine the data, logic dictates that you have to create the data the same way, you have to have the same questionnaires, for example.
While ‘logical’ to the majority, one awardee did not agree with a multicenter design. The fifth PI, epidemiologist Dr. Winkelstein of the San Francisco site, protested vociferously at the first meeting of the awardees. Immediately following that meeting, in a letter addressed to the NIH program officer, Winkelstein argued to retain the original uncoordinated and individual project approach outlined in the initial RFP. His reasons were scientific, logistical, and budgetary. For instance, as Winkelstein (1999) stated, ‘the overriding argument that I made was that when you don’t know very much about a situation such as this epidemic, you’re better off with a diversity of approaches’. In his view, the cross-site standardization demanded by a multicenter study limited the investigation, making it less likely to find the cause of AIDS. While the San Francisco site helped in creating some of the MACS protocols, Winkelstein never joined the MACS and, as the next section shows, the SFMHS diverged in significant ways.

The data and specimens

Initial negotiations were all conducted under the shadow of the study’s central challenge: how to investigate a disease with an unknown etiology? Generally, the biomedical community already leaned in the direction of assuming an infectious agent:

We had absolutely no idea what, how big it was going to be or how it was transmitted or anything. [Quickly] we realized that it was not only gay men … but hemophiliacs, and there were starting to be reports of blood transfusion, it was pretty clear that not only was it an infectious agent but it was an agent that tended to be transmitted sexually, by blood, by and with patterns that were familiar to us with Hepatitis B, for example.8

But other explanations for AIDS were proliferating in scientific circles (Epstein, 1996). Was it the amyl nitrate inhalant (‘poppers’) popular in gay bars and bathhouses? Was it an outcome of gay sex itself: semen in the rectum acting as an ‘immunosuppressant’? A ‘multifactorial hypothesis’ posited an exhaustion of the immune system in the combined face of drinking, illnesses, sex, and drugs. Or was it the Kaposi’s sarcoma cancers or CMV that had initially drawn attention to the disease? A broader set of explanations than an infectious agent had to be considered in the study.

As such, the strategy was to prepare simultaneously for different possible disease mechanisms, whether these were rooted in ‘population and behavior’, ‘systemic pathology’, or ‘microscopic’ causes:

I think one of the smartest decisions that we made — it’s always nice to look back and realize you did do something right at the beginning — we decided to collect specimens and set up repositories, and that was an extraordinarily wise decision. So initially we collected everything they [the participants] could possibly give us, including blood, semen, urine, feces, and … saliva.9

The list goes on, including demographic and psychosocial surveys, throat swabs and gargles, as well as anal pap smears. As one member put it, ‘we were ready to handle just about any cause, as long as it wasn’t aliens’.
Specimen archives stored preserved aliquots, which were made available to participating researchers:

We sent half of the preserved specimens to a national repository in ... Bethesda. That all had to be standardized. In other words, how do you freeze cells so that they are useful? And all that sort of thing. So it took a while to get that all worked out. But we got it done in six months, so it went pretty well.10

PIs report that the discussion of how to collect, prepare, and preserve biological materials was technical but largely uncontroversial; this was the specialization of specimen archivists at NIH (Radin, 2013).

The instruments

Building the kernel’s third component involved the creation, procurement, and standardization of instruments. Instruments are the inscription devices that translate participant’s behaviors, histories, and bodies into data or specimens. For example, flow cytometers were used to measure T-cell subsets, crucial to the regulation of the immune system. In the early years, this measure was one of the only ways of prospectively detecting pathology in advance of AIDS-related symptoms. The MACS used ‘identical flow cytometers, monoclonal antibodies, and analytic procedures’ (Giorgi et al., 1990: 173). Each site purchased the same make and model of the flow cytometry machine, reagents were acquired from the same vendors, and each machine was calibrated using the same brand of ‘alignment beads’: tiny glass globules that were more reliably regular than any human blood cell (Keating and Cambrosio, 1998).

In tandem with specimen collection, MACS-wide questionnaires generated behavioral data and medical histories: sexual practices and partners, drug use and medications, places frequented, occupation, family history of disease, and so on. The PIs debated the length and content of the baseline questionnaire (e.g. what drugs, what sexual practices), but in the end, all agreed that it was long but thorough, taking approximately 45 minutes for each participant to complete.

The participants

The fourth component of the MACS kernel is the participating men. Members report that it was challenging to determine how to recruit gay and bisexual men11 willing to donate their time and bodies to this investigation, in a general environment of prejudice toward homosexuality and growing fear of AIDS. It was, however, also a time of community building among gay people, of awareness raising and mounting activism about AIDS.

One of NIH’s award criteria was a demonstrated access to gay and bisexual populations for subject recruitment. The MACS recruitment focused on men ‘at risk’ of contracting AIDS – a nebulous designation used by epidemiologists and referring to segments of the population identified as more likely to contract the disease (Oppenheimer, 1992). Before the awards, all the sites had forged connections with medical clinics, community leaders, and establishments such as community centers,
gay bars, and bathhouses. MACS’ four centers created a roughly common recruitment protocol, which made comparing findings across sites methodologically easier. They decided to ‘recruit gay men from wherever they were willing to come and without regard to where they lived or without regard to any sort of definable population’. However, using a convenience sample limited their ability to generalize to the broader gay population, or to any other population. For the MACS, a consequence was a common phrase appearing in almost all their early publications; the quote below is drawn from a study of the Chicago MACS, noting a challenge to generalization across the city’s gay men, and to the study as a whole:

although it is not possible to assume that participants are representative of all homosexual men in Chicago, their sociodemographic characteristics suggest that they resemble both homosexual men diagnosed with AIDS in Chicago and cohorts being studies in other metropolitan areas. (Emmons et al., 1986: 334)

Inclusion and exclusion in founding the kernel: ‘at risk’ cohorts versus population sampling

A debate that occurred during the participant recruitment protocol design serves to illustrate the ‘archival’ qualities of the kernel, the constitution of which involves practices of remembering and forgetting (Bowker and Star, 1999). How a kernel is assembled enables, but also forecloses, the examination of particular objects of research. The kernel is ‘ontologically oriented’, that is, it is considered for its capacity to support the study of emergent and forecasted objects. However, as an archive it is also appraised for its parsimony, as every addition is also a long-term commitment (in labor, in time, and in funds). In this sense, the kernel supports multiple ontologies (Mol, 2002), but certainly not every ontology. Exclusions may simultaneously strengthen the ability to investigate certain objects, while barring others – this was the case with recruiting heterosexual men for the SFMHS but not in the MACS.

As noted above, in addition to the four awards that led to the MACS, NIH also awarded an additional site in San Francisco that remained an independent investigation: The SFMHS. The SFMHS varied in several ways from the MACS, but most notably in its recruitment protocol. With no test available for the disease at the time, the MACS had decided to recruit gay and bisexual men who were at risk. They targeted this ‘at risk’ population because it seemed more likely that some participants would eventually develop AIDS. In contrast, the SFMHS’ approach was to ‘study a geographically confined area and literally do a sampling of households, individuals in the Castro District and in other sections where there were known to be high concentration of gay men’. This sampling protocol enabled SFMHS to generate a cohort that was representative of a regional population, with its attendant powers for generalization.

However, from the perspective of the MACS, it was felt that the SFMHS population sample would include many fewer participants who would come to have AIDS. In particular, by sampling a population, the San Francisco study had recruited heterosexual men, the only award to do so:
Eventually we were one of the few places in the country which could give any information on sexual practices of heterosexual men because we had all this information equally from heterosexuals and from homosexuals. (Winkelstein, 1999)

The MACS did not have these subjects and their derived specimens and data, but it was believed that the MACS recruitment strategy increased the likelihood that their cohort would include men with the syndrome. The MACS PIs believed this carried more weight than producing findings that were generalizable to a broader population. That is, the expectation, or rather the crucial goal, for the MACS was to maximize the ability to target its central object of research: AIDS.

In 1985, 2 years after SFMHS’ founding, the study tested their cohort for the virus and found that none of the heterosexual male participants were HIV-positive. The SFMHS PI, Dr. Warren Winkelstein, saw this as a crucial finding: ‘That was terribly important. We didn’t know how this disease was being transmitted. I mean, we had some ideas, but we didn’t know. Maybe there would have been 20 percent infected in the heterosexuals. How would we know?’ (Winkelstein, 1999). Notably, this is not a finding the MACS was capable of generating with its exclusively gay and bisexual cohort. ‘Heterosexual infections’ is not an object of investigation that can be crafted from the MACS kernel. However, for many in the MACS who were facing what one NIH program officer described as ‘the ascending limb of an epidemic’, Winkelstein’s finding was ‘a luxury’, that is, no men in that subsection of the SFMHS cohort could serve to investigate the natural history of AIDS.

We recount this debate around study design and cohort make-up to draw attention to the founding moments of a kernel, which variably render some features of the world in detailed texture and others not at all. Each inclusion strengthens the ability to make generalizable claims, track phenomena of interest, or investigate particular objects; yet, each exclusion also potentially creates a domain of imperceptibility. Even as it created a domain of imperceptibility, it also enabled a more targeted investigation of the central object of research, AIDS. More generally, then, the objects of investigation that an infrastructure supports are deeply entangled with the set of resources and services that it makes available for repurposing.

**Repurposing the kernel: from AIDS to HIV/AIDS**

We started our recruiting towards April 1st [1984], and … just few weeks later, NIH told us this is coming out in *Science*, and this changed the whole ball game … Now you knew what caused it. (Charles Rinaldo, PhD, PI of Pittsburgh site)

*Repurposing* means reassembling the existing resources and services of the kernel to generate new objects of investigation. In this section, we focus on the period following the discovery of HIV, the first ontological transformation we track. US Secretary of Health and Human Services Margaret Heckler announced HIV as the causal agent of AIDS on 23 April 1984. The discovery also promised a pathway to an HIV test, but this would not be commercially available for nearly a year.
The announcement of HIV came twenty-two days into MACS’ participant recruitment. The study had been funded for six months; the four geographic locations selected; questionnaires developed across the sites; and sampling methodologies standardized (in words and documentation, though not yet in practice). But members were still building the cohort across four cities and had not yet begun what would later become the biannual routines of data and specimen collection.

The discovery of HIV as the cause of AIDS is the defining moment in the disease’s history. From a murky cloud of possible configurations – what may have been a different disease in each of the four cities of the MACS, what may have been environmentally or behaviorally caused, and what may have been non-transmissible – AIDS became viral and infectious.18 The announcement of this new entity centrally addressed one of the founding goals and the primary purpose of the MACS: understanding its cause. One may expect that an ontological change of this magnitude would shake the MACS to its foundations. It did not.

By all accounts, the MACS largely went on as before. Identification of HIV was not the end of the MACS, but, instead, the beginning of a new research program. Following the discovery of HIV, MACS members describe an increased focus on behaviors linked to disease transmission, particularly sexual, and an emphasis on the facilities, instruments, and methods of virology. Overall, though, respondents do not recount being deeply concerned about their study, holding emergency meetings or elaborating methodologies, recruitment strategies, or questionnaires.

The explanation for this is the broad scope of the kernel and the ways in which it was initially assembled: the key specialists for studying this new entity, such as virologists and their associated laboratory techniques, were already available to the MACS; the most likely medium of transmission, blood, was already being collected and stored; and the sexual behavior most closely associated with transmission between men, receptive anal intercourse, was captured in the behavioral questionnaire. In the years to come, these were their objects of investigation. The MACS, at the historical moment of identifying this new causal entity, was strikingly well positioned to study the disease as a matter of epidemiology and as an object of bench science.

Extant data, specimens, instruments, and experts were repurposed to investigate new research objects. Those same data originally collected to understand whether AIDS was multifactorial – say, the meeting point of alcohol, poppers, and semen – later were used to understand HIV’s natural history: not ‘does alcohol cause the disease?’ but ‘does alcohol accelerate decline of the immune system?’ (Kaslow et al., 1989).

If anything, they were collecting too much information and materials! And so in the proximate period following the announcement of HIV, we find no new data or specimen collection, and, over time, the specimen collections were significantly reduced; what we call shedding:

We stopped taking stools. We stopped taking urines. We stopped taking semen. We stopped taking throat washings. All we took then is blood. Blood, different amounts of blood and so forth, different blood products.19

Shedding was not immediate, or uncontroversial. With fewer materials available, each cut could potentially make future studies more difficult or impossible. Many MACS
members recount frustration during the debates that led to reductions in specimen collection.

With HIV, many of the previously posited causal explanations for AIDS, such as the ‘multifactorial hypothesis’, largely faded from scientific interest. But shedding from the kernel led to a more determinate outcome: some of these causal explanations could no longer be explored at all. Most notably, with the reduction in specimen collections, the multiple ontologies of causation that the MACS had initially been designed to research, simultaneously ceased to be of interest and the kernel irreversibly ceased to support their investigation.

As Bowker and Star (1999) note in their study of long-term collections and classifications:

> the decision not to collect is the most difficult to take for people maintaining any sort of collection … There are always practical budget and storage issues. These are balanced against two other factors, the need for a well-ordered and in some sense parsimonious repository that can be used, and the side bets that are made about what material will be useful in the future. (p. 69)

Not all specimens deemed irrelevant to the immediate study were shed from the kernel, and those that were, were not shed evenly across each of the four sites. As the Los Angeles PI describes: ‘Here, at UCLA, we did continue to collect semen specimens … longer than the other centers did’. Such site-based diversity of collection is primarily due to the ongoing individual interests of MACS researchers on topics other than HIV disease. Each senior member had come from other research topics before AIDS, such as CMV or hepatitis B; MACS’ vast base of subjects, collected human materials, and thorough inventory of behavioral data could be used to resurface those objects.

This is what Howard Becker (1960) has called ‘side bets’. If the ‘main bet’ of the MACS was the investigation of etiology and transmission, side bets were all those additional objects of research that members ensured could be investigated by the purposeful design of the kernel. Primarily dedicated to the study of AIDS, the kernel also embodies multiple studies only orthogonally linked to that disease. Over time, many MACS researchers resurfaced interests that had led them to AIDS research in the first place:

> We’re still going back. I have students, postdocs, we’re still going back … into those freezers of 25 years ago and getting tremendous benefit out of those specimens that they were so gracious to give us, by the men throughout those years.

The long-term, highly controlled specimen archives, coupled with detailed questionnaires and ongoing possibilities for adding new questions or specimen collections, make the MACS an invaluable resource to explore many medical, biological, and sociological phenomena. Such side bets add flexibility in that they enable new kinds of investigations even if they are outside formally articulated goals.

MACS resources have been repurposed hundreds of times to investigate old, new, and recurrent objects. New objects of research are ‘made’ from the set of resources and services. Each resource can be approached anew or combined in novel ways, in order to tackle emergent or recurrent objects. Resources are also on their own trajectories of change, that is, every year a new batch of specimens and data provides longer views on
the disease, and each year the disciplinary experts bring to bear novel developments in their fields. The next two sections explore changes to instruments and the participants themselves, resulting in altogether new objects of research; in both circumstances, the kernel’s resources facilitated their investigation.

**Elaborating the kernel: from an ‘at risk’ cohort to serostatus groups**

Changes to the kernel *elaborate* if they afford additional depth and detail to extant objects. Knowledge terrains loosely defined in the past are filled out using new instruments and categorizations. In the process, refined phenomena and newly subcategorized objects are investigated at a more detailed granularity.

In a second instance of ontological transformation, we examine the period in which the cohort of ‘at risk’ men became either HIV-negative or HIV-positive (known more formally as serostatus). This elaboration is centered on the development of the HIV test: not a single approach or a circumscribed event, but a series of techniques and technologies investigated, applied, and standardized over the biography of the MACS. Following this elaboration, objects of investigation multiplied to include, most centrally, the newly diagnosed seropositive men, but also, for instance, the consequence of these diagnoses on their behaviors, identities, or self-care. Ian Hacking (2000) has called such categories *interactive kinds*:

> classifications that, when known by people or by those around them, and put to work in institutions, change the ways in which individuals experience themselves – and may even lead people to evolve their feelings and behaviors in part because they are so classified. (p. 104)

Additionally, and importantly for understanding any change to the kernel, the HIV test, as a new instrument, itself became an object of study for the MACS. In order to enact serostatus across the four geographical sites, thousands of participants, and a prospective future, the HIV test needed to be evaluated for the consistency of its results and execution.

In shifting the cohort from ‘at risk’ to HIV-positive or HIV-negative, the kernel was transformed in ways that proved critical to the MACS study design. HIV-negative men were not shed from the study, rather they came to stand in for different phenomena of interest. For example, HIV-negative men could at times serve as controls in studies seeking to compare HIV-positive and HIV-negative conditions, thus revealing, by comparison, the natural history of HIV disease. But, these same men could potentially also serve to continue the investigation of old objects, as they were still ‘at risk’.

**Testing the test: the HIV test as an object and enactor of ontological change**

In order to understand elaboration of the cohort by serostatus, we first need to take one step back from the test, tracking the introduction of this new instrument into the kernel. MACS researchers participated in evaluating the multiple testing kits that started to
become available about a year after the formal announcement of HIV in 1984. The methods of HIV’s discovery had almost immediately led to the commercial production of detection kits, initially designed for testing transfusion blood supplies. Kits are commercial packages for conducting specific assays, marketed as ‘neat, time-saving and error-free alternatives to the messy work of measuring out raw ingredients’ (Jordan and Lynch, 1998: 786). By 1987, about 10 companies were offering kits.22

MACS researchers trialed these kits for diagnostic use with human subjects. Efforts focused on defining a single testing protocol that provided consistent results and could be conducted with ease across the four MACS sites repeatedly over time. Very quickly the laboratory community had determined that the most accessible of these tests used the enzyme-linked immunosorbent assay (ELISA) technique, and so, the focus of MACS research was to determine the most reliable and easily interpreted version of ELISA kits. That is, one set of considerations in testing the tests was the sensitivity and specificity of ELISA kits, but another set of considerations, more pragmatic, was usability within this multi-sited longitudinal study of thousands of men.

Within a few months of starting to test the test, another approach using the Western Blot technique revealed the many false positives and negatives ELISA methods produced (Nishanian et al., 1987). In the MACS (and more broadly), it became a common practice to follow up all positive ELISA tests with a Western Blot. Even with this confirmation method, several members report a low, and unreliable quality to early testing materials: ‘these weren’t all the best tests when they first came out … There were many months that we worked very hard in the different MACS sites, and the laboratories worked very hard to decide on which test we would use’.23

Such problems with the test, and ongoing changes to testing protocols in the years to come, presented immediate diagnostic problems, but for a prospective study they also posed a challenge to the comparability of MACS’ longitudinal data. A member of the MACS coordination center recounts the issue from a data standpoint with the creation of a data file called ‘HIVDef89’, or ‘definition of HIV serostatus in 1989’. Since the specific outcome of HIV serostatus (i.e. positive, negative, or indeterminate) varied by testing kit, by protocol, and by criteria for markers, the ‘HIVDef’ files permitted a commensuration across definitions, enabling MACS investigators to easily convert historical data on subjects’ HIV status regardless of the diagnostic test used at the time:

we worked with the investigators and the labs and came up with different algorithms for what is a seropositive, what is a seroconverter … [The result was the creation of] … a file that we call ‘HIVDef89’.24

‘HIVDef89’, and ‘HIVDef87’, ‘HIVdef86’, and so on, allowed MACS investigators to use past operationalizations of ‘HIV-positive’ without having to ‘redevelop who’s HIV-positive, who’s HIV-negative, [or] when they seroconverted’.25 That is, the commensuration of HIV definitions across these data files helps facilitate analysis to standardize the use of old findings for new research, or repurposing longitudinal data for contemporary objects of investigation. This process of capturing ‘HIV-positive’ in the data – a trajectory starting from a participant, to his blood, to a test (or multiple tests) of that serum, to the codes in a file that stand-in for serostatus – reveals the complex entanglement of the
activities that sustain availability for use and reuse, and the kernel’s resources themselves (Ribes, 2014).

Adding a new instrument to the kernel involves a great deal of subtended work that will cascade across the operations of a research infrastructure and into its future. While, in general, the trajectory of the test can be described as establishing cross-site and longitudinal comparability of serostatus, in actuality, what are often taken as clear categories – HIV-negative and positive – are worked over again and again, and throughout the history of the MACS. Enacting ontological distinctions for each specimen and participant across the four sites year after year, is work done and sustained at the level of the kernel. After 1985, HIV-positive and HIV-negative became crucial categories for all AIDS research; for individuals drawing on MACS resources, these categories were delivered (as data) to investigators largely unaware of the enormous and ongoing backgrounded work of cross-site and long-term coordination.

**Recurrent objects: from a fatal to chronic disease**

Our final vignette focuses on yet another ontological change to MACS’ research objects, namely, the consequences of highly active antiretroviral therapy (HAART). HAART is the combination drug therapy that, at its best, virtually halts the replication of the HIV virus in the body, reducing viral load to undetectable levels, and categorically transforming the disease from fatal to chronic.26 The event most often associated with these treatments is the 1996 International AIDS Conference in Vancouver. Treatment came to divide the history of AIDS into distinct ‘pre and post-HAART eras’. For many MACS participants with access to HAART, treatment marked a transition from ‘dying from AIDS’ to ‘living with HIV’. For the MACS investigators, their objects of investigation manifested in new ways and at differing timescales.

In a scientific research trajectory, Gaston Bachelard has described such objects as characterized by *recurrence* (Rheinberger, 2000, 2005). Rather than following a linear trajectory of constitution, investigation, and comprehension, these objects ‘recur’, displaying recalcitrance to understanding, appearing in new but connected forms, still the same object but now shifted through relations of observation, intervention, or mutation. HIV disease is such an object, first challenging a tidy understanding of ‘a natural history of AIDS’, and then, through medical intervention, splintering into multiple ‘treated histories of HIV’. A relational and historical ontology is needed to make sense of such shifts: relational, in that these objects are tied to many investigators, instruments, and interventions, and historical, in that these relations have been recast over time.

With the advent of HAART and an emerging discourse around the management of a now chronic HIV disease, why continue the MACS? Rheinberger (2010) notes that scientific objects ‘are distinguished precisely by a certain lack of knowledge: they are relevant to research only as long as they leave something to be desired’ (p. 8). Similarly, a research infrastructure is relevant only to the extent that it offers exclusive, or more compelling, access to ongoing and new objects of investigation. The NIH, MACS’ core funding bodies, presented these same concerns from an institutional perspective:

MACS was threatened with extinction several times … There was a question at NIH as to whether or not the MACS outlived its usefulness. What else were we going to learn? Particularly
after treatment became available … That was a point at which they said, ‘well this is a natural history study, and the natural history of AIDS is over if we are going to start treating all these people’ … That was the kind of tenor of the conversation.27

It is their future-oriented quality that makes objects of research. Once they lose their ability to surprise – their recalcitrance to prediction and control – they soon fall out of favor as scientific objects. This does not mean that they cease to be, but rather that they are ‘shelved’ as understood phenomena in the world. Without ongoing objects of research, the MACS would shift from being a research infrastructure to a relic.

The MACS, as a whole, is continuously engaged in recasting itself as a study that has not fallen into this realm; it continues to extend its research into wholly new objects of study, while simultaneously sustaining and regenerating its core resources. MACS’ senior investigators have been at pains, at each of its five-year grant renewals, to demonstrate this to their funding institutions and their respective research communities (cf. Patel, 2012): to demonstrate that MACS’ objects will continue to broaden and deepen, to show relevance to new fields of inquiry, and to link to new phenomena.

Extending the kernel: treatment, chronic illness, and aging

As with the two previous ontological shifts, the development of HAART treatment was not an event, but a protracted series of trials and investigations. For the MACS, it necessitated an extension into uncharted terrains of research, drawing on new specialties and techniques. Extension refers to changes to the kernel that enable investigations of entirely new research objects. MACS, in the post-HAART era, certainly repurposed existing data, specimens, instruments, and experts, but as a chronic disease HIV presented newly relevant objects that MACS scientists had never systematically investigated. To do so, they solicited new specialists and their associated techniques and instruments. They began collecting new specimens and questionnaire data, and they began asking new research questions.

HAART presents its own unique life difficulties for those taking the treatment. Rather than a natural history of AIDS, today, the MACS positions itself as a study of the treated histories of HIV disease. This is not a single treated history: a chronic disease is tied closely to individuals, to their particular manifestations of the disease, and to their own medical, psychological, and life issues. Plus, the treatment itself is a source of problems: ‘These drugs are miracles but they are not curing the men. And they still have detrimental [effects] to their health’.28

Treatment regimens are themselves in motion. Regimens for each person are tailored to maximize effectiveness and to minimize side effects, and adjusted as new drugs become available. In its early years, the HAART ‘cocktail’ required complex dosage schedules, demanding that dozens of pills be taken over the course of a day, some with food, others without. Today, some of the most effective treatments, now often just called antiretroviral therapy (ART), require only one or two pills containing multiple pharmaceuticals. But most HIV-positive men in the MACS have been on many different regimens, traversing generations of the treatment over the years. Few men from the study can stand in for the long-term effects of one single treatment.

The participants of the study, too, are in motion. As is commonly stated, they went from ‘dying from’ AIDS to ‘living with’ HIV, but this phrase veils the challenges of
individual disease trajectories. Post-HAART, the most fortunate men were less ill, and this meant they were living new lives. However, these men, thousands from the original cohort, were aging. ‘As the men aged, that has a tremendous effect on their immunity and the way they handle this virus. Or does it? And, that’s the question’.29

The 2008 NIH call for proposals to renew the MACS extended the studies’ core research questions into these domains:

These awards will support ongoing MACS research projects and the initiation of new projects on … the effects of long-term HIV infection and treatment on the aging process and the impact of HIV therapy on cardiovascular disease, liver disease, renal disease, neurocognitive impairment, cancer and other outcomes. (NIH, 2008)

Many of the old objects of study remained – for example, time from infection to AIDS; patterns of transmission – but at the nexus of HIV disease, treatment, and aging, a whole new set of objects emerged that the MACS had never investigated and which required new approaches and new investigators:

The initial thing [AIDS] was obviously a malignant disease, and so we recruited investigators who were interested in malignant disease. Then, it became apparent that there were metabolic complications of treatment, so we got people who were interested in metabolism to work with us. And then it became apparent that a lot of these men had chronic hepatitis and they were living long enough to have problems with liver disease because they had viral infections of their liver. So we got people interested in hepatitis to … work with us. And so that’s how it went: the growth of the investigator pool.30

Some of these new researchers are able to assemble their objects from the existing kernel. But new objects, such as chronic illness, drug adherence, and aging, required novel data and specimens. For example, beginning in 1998, a new questionnaire was administered investigating medication adherence, and in 2011 a questionnaire began generating data on ‘instrumental activities of daily living’, such as the ability to perform housework, get groceries, or manage finances. These new data made possible the assembly of particular objects related to the study of chronic illness, treatment, and aging with HIV.

As co-morbidities became of increasing interest in the late 1990s and early 2000s, some of the specimens initially shed following the discovery of HIV were collected once again. For example, anal pap smears returned to the collection routine, not for detecting the etiology of AIDS (their initial purpose), but for the study of cancers or stomach flora. While there is a twenty-year gap in the archive, specimens collected in the mid-80s were never shed. Recently, they provided a comparative window into, for example, a pre-treatment-era incidence of anal cancer. It is only through the extension of resources and services that certain new objects of investigation – like medication use, aging, and co-morbidities – can be assembled from the kernel.

Facilitating extension

New investigators targeting new research objects usually find that not everything needed is ready at hand. The kernel only sometimes can be used ‘as is’ to take on new research
objects using existing data, samples, participants, experts, and instruments – what we have called repurposing – and so the MACS also reflexively facilitates change to its kernel.

Approximately at the time of HAART (circa 1996), the MACS began to refer to itself as ‘an infrastructure’. In a formulation nearly paralleling our own definition of the kernel, the 1998 NIH RFP to renew the awards defined the MACS as

The MACS research infrastructure offers a unique focus for study of HIV pathogenesis including:

1) A fifteen year prospective cohort study …
2) A repository of clinical specimens of serum, plasma, …
3) Extensive clinical, epidemiologic, and statistical expertise …

(NIH, 1998)

Calling the MACS an infrastructure emphasizes the utility of its resources for a broad community of researchers, undoubtedly helping in securing renewals. However, it also speaks to what we have observed as a self-reflexive effort to make the kernel available for the investigation of new objects and, particularly, to facilitate the kernel’s own transformation by the addition of new resources.

The Study Specific Concept Sheet is an example of such an effort to facilitate change to the kernel (MACS, 2011). All new investigators drawing on MACS resources are required to file this form. It is a six-page document requesting basic information about proposed study objects and methods, the investigators, and what will be required from the MACS. In principle, each of the four components of the kernel is available for repurposing, elaboration, or extension, though use of any particular resource carries its own unique burden. For example, data, particularly recent data, are valuable within ongoing core MACS studies and so there is often a delayed release schedule before they are shared with outside collaborators, giving core members a privileged window of time to conduct their research.

The Concept Sheet routinizes new membership and new objects. It is a form-based obligatory passage point for gaining access to MACS specimens, data, and participants. The MACS has staff that facilitate access to and understanding of these resources:

We’ll work with [new investigations], establish who will be included in the study, and what specimens will be selected, from what visits … We work with the repository, and submit a request to the repository to have the samples sent to the investigator.31

Coordination center staff provide this service, acting as facilitators between a vast archive and the specific materials needed for an individual study.

When we approached the MACS to conduct this sociological study of ‘research infrastructure’, no one blinked an eye at our terminology or choice of research object. By completing a Concept Sheet, our own study was granted access to the MACS (its data, experts, and participants in Baltimore/DC). We, the first archival and ethnographic study of the MACS, became a part of the study, in the sense that we gained access to some of its resources and are managed similarly to other investigators. On an annual cycle, we continue to receive requests from MACS staff for updates on our research, funding,
findings, and publications. These are in turn added to a database by MACS staff, who track and record ongoing studies.

**Inclusions and exclusions**

Our focus in this article has been to track the set of resources and services offered as part of the kernel of a research infrastructure. The kernel, as a constellation of concepts, seeks to explain the ability of a research infrastructure to prepare for and adapt to ontological changes, a capacity we call technoscientific flexibility, and which we have explored concretely through processes of repurposing, elaboration, and extension. But throughout the article, we have also pointed to systematic limits in the flexibility of any kernel: not all objects of investigation can be assembled from the materials made available. Considered as an archive, the kernel supports certain investigations of a past, present, and future. But all archives comprise activities of selection and clearance of its materials, and practices of curation preserve some relations and not others: in sum, creating domains of imperceptibility. Elaboration and extension of the kernel cannot always shift those domains, but, paradoxically, we also find that the exclusion of some objects may strengthen the ability to investigate others.

We have discussed how the MACS focused on sexually active gay and bisexual men in the United States at risk of contracting AIDS in 1983–1984. MACS did not, and still does not, include heterosexual men within its cohorts. Objects of research specific to those populations cannot be assembled from the MACS kernel. But, conversely, at a time when an agent was undetermined, prevalence rates unknown, and a test unavailable, the decision to focus on men from an ‘at risk’ population was felt to strengthen the likelihood of recruiting subjects who would develop AIDS.

We have discussed how other materials, initially part of the archive, were strategically shed for a window of time, such as anal pap smears during the 1990s. Contemporary investigations of anal cancers within the MACS rely on archived and newly collected smears, but suffer from the gap left by shedding this collection. Certain research objects can be crafted from that archive, and others cannot; for instance, studying the interactions of anal cancer and early HIV treatments in the mid-1990s is challenging because it occurred during a period when the relevant materials were not collected MACS-wide.

Meanwhile, other features of the kernel have been actively extended through recruitment efforts that target underrepresented populations in the initial cohort. For example, recruitment of additional men focused on diversifying race and ethnicity. The first cohort had been approximately 95% White, the first recruitment protocol having favored White men. Recruitment in 2001 led to an additional 1356 men, of which 72% were non-White. Such efforts are reflexive attempts by MACS members to extend to new topics that they came to recognize as important or excluded, but remained essentially impossible to study in the first cohort alone. However, the archive never stretches back before the addition of new features; thus, comparative studies of race and ethnicity are more difficult to conduct with data and specimens collected before the addition of the second and third cohorts.

Even with such extensions, there are objects of research that the kernel systematically does not support and that have never been the MACS’ goal to incorporate. It will always
be impossible to say something specific, from within the MACS kernel, about the variation of HIV disease in non-Americans or women. Critical scholar Paula Treichler (1987) famously asserted that medical studies of AIDS were founded on men’s bodies, leading to years of confusion and misinformation about HIV transmission and disease in women. In 1993 and partly in response to such criticisms (cf. Epstein, 2007), the NIH founded the Women’s Interagency HIV Study (WIHS). The WIHS was partially informed by the MACS, often referred to as MACS’ ‘sister site’: They share some geographic centers, many protocols and instruments, and even organizational members. Such a forking of kernels facilitated commensuration of data and comparative analyses, while providing a needed autonomy for each study to investigate distinct objects. However, the scientific horizon of exploration afforded by a 10-year gap in the archive can never be recovered.

MACS supports the investigation of a broad (and emergent) range of objects by regenerating access to a set of resources and services. What can be investigated, and how, is deeply tied to those materials, data, and the sources of collection – a bounded technoscientific flexibility. The objects of investigation that a research infrastructure will support are deeply entangled with the constitution of its kernel, a changing, but not fully protean, set of resources and services.

**Conclusion**

We have sought to distinguish the concept of ‘emergence’ from ‘de novo’ by tracing the practical links between investigations of novel and recurrent research objects on the one hand, and curated archives, routinized specimen collections, and calibrated instruments, on the other.

‘Emergence’ draws attention to the surprising or recalcitrant qualities and relations of objects of investigation, qualities and relations that could not be fully outlined in advance of their practical investigation. But surprises need not be altogether unrelated to what came before, or ‘de novo’. In order to help think through this relationship, we have drawn on two felicitous phrases: ‘organized chaos is the heartland of innovation’ and ‘chance favors the well prepared’. The kernel resides at this intersection, characterized by efforts to accountably routinize and standardize its features, but also valued as the combinatorial components for innovation.

On its own, the kernel is not sufficient to investigate new objects. Rather, the kernel provides resources to be strategically assembled by investigators as they craft their experimental systems. Investigators do so by bringing to bear their own skills, research funds, labs, or equipment in coordination with the kernel, selectively drawing from its set of resources and services. Rather than the source or determinant of innovation, the kernel is what organizational members work over to make it available as infrastructure, thereafter to be strategically assembled by investigators to refine phenomena and to craft researchable objects.

The kernel renders certain objects easily inspected; others are a greater challenge, and still others cannot be inspected at all. Stewarding of selected resources and services enables the enactment of specific ontologies and not others – a bounded technoscientific flexibility sustaining a link between emergence and the materials of that emergence. In contrast, something that is altogether novel, completely unrelated to those things that
came before, and disconnected from all practice and routine, will be useless and beyond understanding.

Acknowledgements

We thank Jennifer Feldman for her work during the research, and Steve Epstein and Geoffrey Bowker for insightful conversations and feedback on drafts. We would also like to thank the MACS PIs, staff, and participants for access and insights during this research, especially Michael Plankey, who first introduced us to the MACS.

Funding

This research has been supported by NSF/VOSS award #1322275 ‘Flexible Research Infrastructure: A Comparative Study’ and the NSF/STS Scholar award #1431219 ‘Examining the Technoscientific Flexibility of Research Infrastructure’.

Notes

2. Mirroring Multicenter AIDS Cohort Study (MACS)’s terminology, we use the term ‘participants’ to refer to the men in the study. We use the term ‘members’ to refer to principal investigators, regular collaborators, and staff formally affiliated with the MACS. For quotes drawn from interviews, we have asked if we may attribute the quote; the majority have agreed, but some respondents have asked that some quotes remain anonymous.
3. Interview with John P Phair, MD, Chair of MACS Executive Committee and PI of Chicago site (4 January 2012).
4. Interview with Roger Detels, MD, PI of the Los Angeles site (4 January 2012).
5. The original PIs were David Ostrow, MD/PhD, of the Northwestern School of Medicine in Chicago (a position quickly taken up by John P. Phair, MD in 1985); B. Frank Polk, MD, of the Johns Hopkins School of Hygiene and Public Health in Baltimore; Charles R. Rinaldo, PhD, of the University of Pittsburgh School of Public Health; and Roger Detels, MD, of the UCLA School of Public Health.
6. Interview with Roger Detels, MD, PI of the Los Angeles site (4 January 2012).
7. Interview with Charles R. Rinaldo, PhD, PI of the Pittsburgh site (12 January 2012).
8. Interview with former NIH/NIAID Program officer (18 January 2012).
10. Interview with John P. Phair, MD, Chair of MACS Executive Committee and PI of Chicago site (4 January 2012).
11. Roughly speaking, over time MACS publications have shifted from using identity-based categories (e.g. gay and bisexual) to sexual behavior categories (i.e. men who have sex with men – MSM) – even though the two categories are by no means coextensive, and identification with identities or behaviors is a complex matter (Martucci, 2010). However, the pathway is by no means linear: from the very first MACS questionnaires, sexuality was operationalized as identity based (e.g. a seven-point scale from homosexual to heterosexual) and behaviorally instrumented through structured inquiries into sexual activities.
12. Interview with former NIH/NIAID Program officer (18 January 2012).
13. Interview with former NIH/NIAID Program officer (18 January 2012).
15. Interview with Charles R. Rinaldo, PhD, PI of the Pittsburgh site (12 January 2012).
16. Heckler attributed the retrovirus’ discovery to Robert Gallo and his lab at the NIH, who titled the virus Human T-cell Lymphotropic Virus III (HTLV-III). Famously, ‘because of a cold,
she lost her voice’ (Harden, 2012: 64), and so Heckler failed to credit French scientists Luc Montagnier, Françoise Barré-Sinoussi, and their team at the Pasteur Institute, who first isolated the virus, and called it ‘lymphadenopathy associated virus’ (LAV). HIV was the name given to the virus in 1986, once the multiple attribution claims had been (relatively) settled. To keep an already complex story more straightforward, we use the term HIV throughout. For ease of reference, we also use ‘HIV disease’ as a general term throughout the article.

17. The discovery of HIV is one of the most famous priority disputes in the history of science (Grmek, 1990). Using this date is arbitrary, but not particularly important for the purposes of this article. Before this event, MACS members continued to consider alternative causal explanations for AIDS; shortly after it, they did so almost not at all.

18. HIV as the cause of AIDS was not to remain uncontested (Epstein, 1996; Fujimura and Chou, 1994). While MACS PIs made public announcements in support of the HIV-as-cause thesis, this challenge only peripherally impacted work in the MACS.

19. Interview with Charles R. Rinaldo, PhD, PI of the Pittsburgh site (12 January 2012).
20. Interview with Roger Detels, MD, PI of the Los Angeles site (4 January 2012).
21. Interview with Charles R. Rinaldo, PhD, PI of the Pittsburgh site (12 January 2012).
22. The full history of testing the test is beyond the scope of this article and has yet to be properly accounted for in the literature (but see Harden, 2012).
23. Interview with Charles R. Rinaldo, PhD, PI of the Pittsburgh site (12 January 2012).
24. Interview with member of the Data and Analytical Coordinating Center for the MACS (18 January 2012).
25. Interview with member of the Data and Analytical Coordinating Center for the MACS (18 January 2012).
26. The re-categorization of the disease was complex and debated: as early as 1989, Samuel Broder, head of the National Cancer Institute, declared AIDS a chronic disease (Fee and Fox, 1992). However, the advent of HAART is still seen as a landmark shift in the trajectory of the disease. Our interest is not in pinning down a moment for reclassification of AIDS as chronic, but instead in tracing how MACS responded as the disease changed.
27. Interview with former NIH/NIAID Program officer (18 January 2012).
28. Interview with John P. Phair, MD, Chair of MACS Executive Committee and PI of Chicago site (4 January 2012).
29. Interview with Roger Detels, MD, PI of the Los Angeles site (4 January 2012).
30. Interview with John P. Phair, MD, Chair of MACS Executive Committee and PI of Chicago site (4 January 2012).
31. Interview with member of the Data and Analytical Coordinating Center for the MACS (18 January 2012).

References


Author biographies

David Ribes is assistant professor in the Communication, Culture, and Technology (CCT) Program at Georgetown University. He is a sociologist of science who focuses on the development and sustainability of research infrastructures (i.e. networked information technologies for the support of interdisciplinary science), their relation to long-term changes in the conduct of science, and epistemic transformations in objects of research. His methods are ethnographic, archival, and comparative. See davidribes.com for more.

Jessica Beth Polk is a doctoral student in the Department of Science & Technology Studies (S&TS) at Cornell University. Her research stands at the intersection of the sociology of science (with a focus on medicine), infrastructure studies and systems of data sharing, as well as the study of medical expertise and experience. Prior to obtaining her MA at Georgetown University’s Communication, Culture & Technology Program (CCT), she spent several years working in the healthcare industry.