

# Performing Big Data

## Alan Lomax and The Folk-Song Science

Joe Diebes

At first glance, computers are at odds with live performance. The cold pre-determined logic of the algorithm would seem to negate the spontaneity valued by artists and audiences in even the most scripted and rehearsed productions. My purpose, in structuring most of my operas and audiovisual works around computational concepts, has been to heighten this incongruity into a theatrical tension that speaks to a distinctly contemporary condition. On one hand, algorithmic processes have become increasingly internalized, often determining what people think, do, and want in the world; on the other hand, there is still something—call it a classical sense of free will, or perhaps simply the friction of particular bodies and voices—that resists such automation. This was the kernel of previous works such as *I/O* (2006) and *BOTCH* (2013), where vocalists interact with each other and with machines according to rules cannibalized from computer programs, customer service protocols, and voice recognition devices.

This same tension between agency and technological constraint led me to Alan Lomax as the subject of my most recent opera, *OYSTER*. Though he is popularly known for his recordings of blues and folk music in the 1940s and 1950s, the American folklorist and musicologist spent much of his later career on what we today would call a big data project. In the early 1960s, Lomax embarked on Cantometrics, a system he invented to profile and quantify different styles of folk music from every known culture in the world. It is difficult to overstate the ambition of this enormous institutional project, housed by Columbia University and funded by the federal government, that included an interdisciplinary research team of linguists, anthropologists, statisticians, programmers, and movement analysts—all fueled by the general cybernetic fervor of the time. Harnessing the processing power of an IBM mainframe computer, the goal was no less than to capture a computable image of the entire planet. My libretto for *OYSTER* asks how Lomax, the folk song fellow traveler, could become the engineer of this total

system with imperial dimensions. My vocal score and video design questions whether such a system can make universal claims in the first place.

*OYSTER* premiered at Roulette Intermedium in New York City, and was staged as a public lecture by Lomax (performed by John Rose). From his podium, he explains his “folk song science” in a sung-spoken style, accompanied by live prepared piano and electronics. His presentation is interspersed with a trio of performers (Christina Campanella, Michael Chinworth, and Saori Tsukada) who piece together Lomax’s life, parsing and scanning an extensive, recently declassified FBI file, while performing a choreography of geometric patterns staged by Phil Soltanoff. Tightly woven into the live performance are a series of projected music films that function as audiovisual aids for the lecture, as well as being the musical core of the opera. In each film we see and hear an ensemble of vocalists (the same performers that are live onstage) simulate a song associated with a different region of the world.

## DATA VOCALIZATION

Here it will help to explain the basics of Cantometrics. Imagine a room full of musicology graduate students in 1964, wearing headphones plugged into reel-to-reel tape machines. The actual conversion of individual songs into data was a hybrid human-machine process whereby these trained coders would carefully analyze recorded songs, extracted from hundreds of different local cultures throughout the world. Whether it was a lullaby recorded in an Indonesian village or a choral wedding song from the Balkans, various attributes were numerically rated. Tempo, for example, was rated on a scale from one to five: a quickly paced tune may be rated a four whereas a funeral dirge might be a one or two. Other attributes, many of which had not previously been considered quantifiable, included the degree of vocal harshness, the enunciation of consonants, or how well a group of singers blended together. For each song exactly thirty-seven such attributes were rated, making up its profile.

The contemporary resonance is that Cantometrics was not concerned with individual songs, but with the statistical analysis of thousands of songs to uncover general trends. Enter the IBM360, one of the very first computers available for businesses, governments, and research institutions. Each song profile was encoded onto an IBM punchcard, thousands of which were fed into a computer that sorted and compiled them to find similarities. A vast diversity of local cultures was boiled down to six essential composite profiles, roughly corresponding to the continents. For example, the South America profile was derived from 230 songs that had been recorded in locations as distant as Patagonia, Eastern Brazil, and

Central America. These virtual projections were supposed to reveal the essential pattern beneath all songs of a given region.

The songs for *OYSTER* are data vocalizations of these composite profiles. The idea was simply to reverse the conversion of songs into numbers by converting the numbers back into songs. I was not expecting total accuracy—these were averages not corresponding to particular songs to begin with—but would they bear some resemblance to their source? I treated the cantometric data as my score and followed the thirty-seven ratings as best I could. Some were easily translated into compositional rules that determined things like phrase lengths, melodic range, the amount of repetition, and how the parts related to one another. For example, the profile for Old High Culture (a region extending from Japan to the Middle East) determined that the song must be for a solo voice in free rhythm with melodic phrases five to nine seconds long and a range between a fifth and an octave. It was important for me *not* to do any more than to simply follow these rules. I did not listen to any actual recordings of songs from the region, nor did I do any additional research beyond applying the numbers. I quickly discovered that some fundamental musical aspects were not accounted for at all, such as pitch relationships: I could write a very consonant or a very dissonant piece and still conform to the same profile. Let's just say there was a lot of wiggle room.

Half of the attributes were connected to the performance style, and much of the realization happened in rehearsal with the quartet of vocalists that appear in the films, a group I have worked with consistently for over ten years (Christina Campanella, Michael Chinworth, John Rose, and Saori Tsukada). Through embodying and voicing these ratings it became clear they were loaded with Western notions of technical proficiency, as they involved consciously performing what would be considered poor classical musicianship. We practiced having a very low degree of rhythmic coordination (South America), in contrast to much higher ratings called for by the Insular Pacific profile and, of course, the European one. As for timbre, the profiles called for techniques that are anathema to Western operatic singing such as the ability to produce various degrees of vocal rasp, nasality and constrictedness. Although those characteristics may be found in non-classical voices we were familiar with (Bob Dylan would rate highly in all three), others, such as glottal shake, required consulting the Cantometrics handbook and its description of a "forceful activity in the pharyngeal area" that can be simulated by manipulating the Adam's apple. In grabbing our own throats to conform to the data, we began to have an inkling of the physical bodies that were essentially erased in being encoded into numbers.

I doubt that any of our vocalizations would be mistaken for those actually performed by singers on the source recordings. What they do offer is a way to

hear the totalizing project of Cantometrics itself, and how constructed such a supposedly objective data model can be. The outrageous gap that resulted from a scrupulous adherence to the song profiles was filled in with my own creative choices and the prejudices and ingrained habits that come “naturally” to a group of experimentally minded and relatively educated artists working in New York City. The visual style of the films—carefully framed in a geometrically designed studio environment—further emphasizes that our data vocalizations result more from aesthetic rather than scientific considerations. Each region is color-coded to reinforce the basic operation of encoding the songs into predetermined categories. That the division of six world cultures correlates so easily to a spectrum of the six basic hues (the primary and secondary colors) is a further clue that the system has a certain *a priori* structure. As I mapped Cantometrics to the utterly artificial form of my opera, I began to see the extent to which it was governed by principles of balance, simplicity, and pleasing form. It was a beautiful system—one that I consciously collaborated with as an aesthetic object.

As science, though, Cantometrics was discredited almost immediately by scholars in the newly emerging field of ethnomusicology. The timing was off. Lomax was promoting his system in the late 1960s and early 1970s, just as anthropology was moving away from a clean structuralism, which reduced all societies to a single set of rules and relations, to more localized studies that favored thick description. But if such universalizing projects were at least ideologically rejected in anthropology, Cantometrics was an uncanny precursor to the way music, and almost every other social activity, is subject to algorithmic profiling in the 2020s. The information economy is driven by composite consumer profiles mined by YouTube, Spotify, Netflix, or Amazon that blur any distinction between analyzing what people like and actually constructing their desires. As for data visualization, I am writing this article in the midst of a pandemic where the recurring public service announcement is “flatten the curve.” Is it even conceivable that individuals could have comprehended themselves so statistically, so quantitatively, before the era of cultural computation that Cantometrics helped inaugurate? At a moment where personal choices are often determined by probability, data distribution, and elegant infographics, this seemingly obsolete chapter in the history of musicology can perhaps offer some genealogical clues to the present.

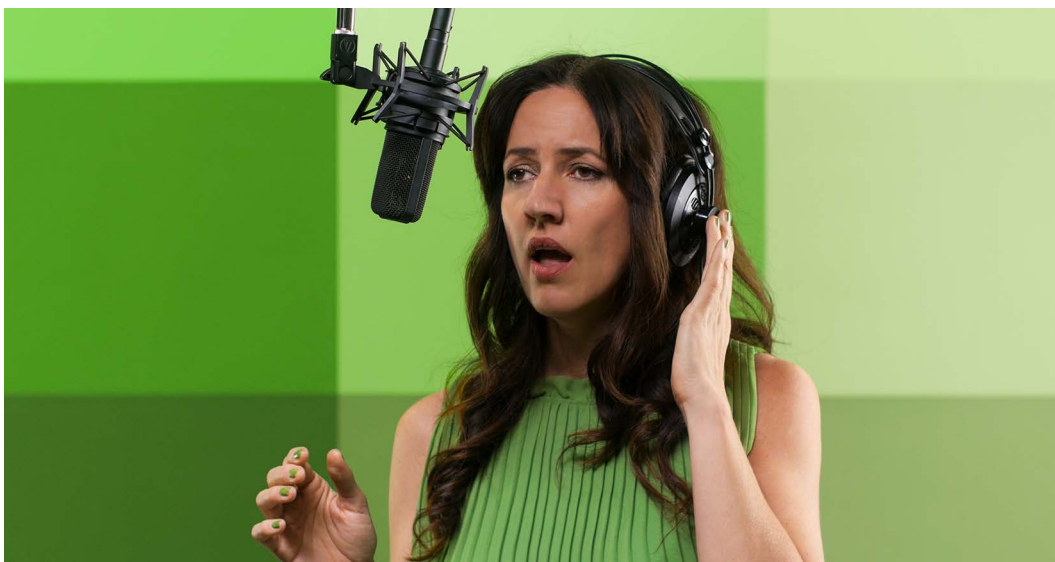
However, the specifically global scope of Cantometrics goes well beyond the construction of domestic subjects to the modeling of *every* culture to conform to the same program. Computers are artifacts of Western European and U.S. military objectives, and this is worth remembering as we consider the proposal that helped secure federal funding for Cantometrics. Folk song style was presented as an indicator of more general social structures already surveyed in such “data



John Rose as Alan Lomax. Damian Calvo, cinematographer (video still).



Data Vocalization. Left to right: Christina Campanella, Michael Chinworth, and Saori Tsukada as IBM360, vocalizing the Africa profile. Damian Calvo, cinematographer (video still).



Christina Campanella as IBM360, vocalizing the Old High Culture profile. Damian Calvo, cinematographer (video still).



*Oyster at Roulette Intermedium*, 2018, New York City. Left to right: John Rose, Christina Campanella, Michael Chinworth, Saori Tsukada, Melinda Faylor (piano). Photo: Steven Schreiber.

banks" as George Peter Murdock's *Ethnographic Atlas*, a list of over a thousand world societies encoded according to a standardized taxonomy of subsistence practices, political organization, gender roles etc. Lomax offered Cantometrics as proof of Murdock's system by identifying correlations between the two studies. Some were quite literal, such as the correlation between group singing and non-hierarchical political formations, that he found in parts of South America; others were less convincing, such as the correlation made between the nasality of vocal quality and sexual prohibitions on pre-marital sex, as he found to be the case in some regions of Spain. In any case, folk songs were a sort of code that needed to be broken. Code-breaking was of course the *raison d'être* of the first computer, designed by Alan Turing only two decades earlier to decode Nazi transmissions during World War II. After the war, the cutting edge of computational development shifted to the United States, a fact that conveniently lines up with a Cold War agenda to gather intelligence not only on the Eastern Block, but the colonies of the Western European empires that were being at least financially transferred to the emerging superpower. That the organizational framework of Cantometrics emerged from such machinations at least begs us to question any scientific claim to neutrality.

## WHO WAS ALAN LOMAX?

This is not yet the makings of an opera. I take as a premise that Cantometrics, viewed from just about any contemporary critical perspective, is thoroughly implicated in the imperial violence that structures the United States' technologically mediated view of the world. But Lomax, the person, is not Cantometrics. In the libretto, I wanted to untangle how geopolitical and technocratic forces are wrapped up in the personal desires and affections of a person trying live his life. Lomax's story is told from two accounts. The first is a recently declassified FBI file that I spent several days engrossed in the Lomax archives at the Library of Congress. Like other left-leaning "folkies" of his time, Lomax was suspected of being a radical communist. He was tracked and harassed for most of his adult life by the federal agency, to the extent to which he was forced to flee the country during the height of the Red Scare and 1950s McCarthyism. This fact alone should complicate any narrative that he consciously became a minion of U.S. imperialism. Excerpts from the four hundred-page FBI dossier are vocalized by a trio of agents intermittently throughout the opera, providing a bizarre biometric analysis of Lomax as himself an object under investigation. His height, weight, passport numbers, visas, parking violations, and just about any quantifiable trace a person leaves in contemporary society are all tabulated and arranged, not unlike a Cantometrics profile.

This turns the ethnographic lens back onto Lomax, making him for a moment an objective specimen under glass. But already it is not that simple. Much of the report consists of secret interrogations that FBI agents conducted with Lomax's acquaintances (another irony is that the FBI called them "informants"—the same term used by Lomax, and anthropologists at time more generally, for locals who help gain access into the culture being studied). We learn, for example, that his apartment neighbors overheard him singing "hillbilly songs" with friends late into the night, that his co-workers regarded him as a "bohemian type" and were bothered by his "slovenly personal appearance," and that he brought undesirables such as "rural folk and minority groups" into the library where he worked. The FBI even trailed him abroad, documenting a trip to Spain in 1951 where Lomax and a woman in her early 30s—herself profiled as being of medium height, slim with blond hair and green eyes—set off on a road trip in search of gypsy songs between Málaga and Granada, winding along roads in a black Citroën filled with recording equipment. I find it fascinating, but also frightening, that such an intentionally forensic document, regulated by bureaucratic protocols, still manages to produce these full-color glimpses into an unconventional life.

But then we hear from Lomax himself, who in the course of his lecture often lapses into anecdotes and philosophical musings. He intimates some of the desires and self-doubts that led a rough-and-tumble folklorist to the helm of such a dominating apparatus. At root, he is a poet. At one point in the lecture, he speaks about his days in the field, his *wanderlust* for song, and how somewhere along the line something changed. The piquant arousal he once felt from a sea shanty or bridal song had become an obsession; he didn't even care about the song he was recording. He just wanted, needed, the next song. Sleepwalking, half a man, he was searching for *the* one that would complete him: the *song-in-itself*. Of course, he never found this Platonic phantasm in the field, but this is where data science and poetry meet, in the dream of the ideal song, the pure virtual song that somehow unlocks a truth beyond the phenomenal world. Like Dante, he is at the midpoint of his life, and after decades of striving and hard experience, he is called upon to map the cosmos. Cantometrics is an ordered vision, even a beatific one, that promises salvation from a murky life of unfinished business. Here, everything is in its place, as it must be in every global database from Cantometrics to Google Earth. Is this poetic vision only an imperialism in another register?

As with many poets, Lomax is haunted by a desire for fame. Before Cantometrics he was never seriously involved with academia, having accumulated America's most extensive archive of folk songs almost single-handedly, and by his own



rules. When he saw the emergence of ethnomusicology, as well as the increasing legitimacy accorded to the human sciences in the age of structural anthropology, he wanted not just a piece of it, but to demonstrate his authority within the halls of the academy. This need for institutional recognition seemed to trump the values he had espoused for decades, particularly his long-held opposition to the homogenizing effect of mid-century mass media on local traditions that he called “the cultural grey out.” He embraced, perhaps with utopian blinders, the ultimate mass media computing machine that established sameness rather than difference across the very cultures whose uniqueness he had defended. But he always remained an outsider. In the penultimate scene of the opera a gang of peer reviewers academically assassinates Lomax (from actual reviews of his 1968 cantometric study *Folk Song Style and Culture*), mainly for reasons I have already enumerated. The tragedy is not just that Lomax had compromised his own values for institutional credibility, but that he had a more intimate knowledge of folk songs in his little finger than the whole bunch of middle-management ethnomusicologists had in their graduate departments.

---

JOE DIEBES combines sound, visual media, and the human voice in various ways. His works include the opera *WOW* (in collaboration with David Levine and Christian Hawkey, BRIC Media Arts), his broken-word opera *BOTCH* (HERE Arts Center), the sound-theatre work *I/O* (in collaboration with Phil Soltanoff, Fusebox/Théâtre Garonne, 2008) and the opera environment *STRANGE BIRDS* (2003). His sound installations, video, and works on paper in have been exhibited at Paul Rodgers/9W Gallery, The Hammer Museum, The Torino Winter Olympics, Yuanfen Gallery (Beijing), Prix Ars Electronica (Linz Austria), and the Liverpool Biennial. In 2018, *OYSTER* premiered at Roulette Intermedium, in New York City.