Lindsay Caplan
Method without Methodology: Data and the Digital Humanities

In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province. In time, those Unconscionable Maps no longer satisfied, and the Cartographers Guilds struck a Map of the Empire whose size was that of the Empire, and which coincided point for point with it. The following Generations, who were not so fond of the Study of Cartography as their Forebears had been, saw that that vast map was Useless, and not without some Pitilessness was it, that they delivered it up to the Inclemencies of Sun and Winters. In the Deserts of the West, still today, there are Tattered Ruins of that Map, inhabited by Animals and Beggars; in all the Land there is no other Relic of the Disciplines of Geography.

– Suárez Miranda, *Viajes de varones prudentes*, Libro IV, Cap. XLV, Lérida, 1658

In this one-paragraph short story by Jorge Luis Borges, "On the Exactitude of Science" (1946), the fictional Suárez Miranda recounts the rise and fall of an imperial project to make a map the same size as the territory it describes. As soon as the awkwardly scaled artifact is complete, however, its prospective users recognize its absurd inadequacy and abandon it to be absorbed back into the ground it was intended to figure.

Borges's image of these threadbare vestiges – the reference to which became something of a postmodern proverb in the second half of the twentieth century – stands as a warning against confusing a thing with its representation. The results are more than impractical; they are dangerously fantastical. It is a fantasy to think we can stand apart from reality and grasp it with the proper, total prosthetic. There is no ontological outside from which our vantage is secure and sacrosanct. Nevertheless, there is today a renewed attempt to conflate the map and the territory. From the NSA's deliberate stockpiling of data and Google's relentless collection of incidental personal archives like old emails, Facebook posts, and website cookies, "Big Data" is information amassed to the point of incalculability. Not quite map and not quite territory, these archives are as vast and unwieldy as the phenomena they seek to chart and define.

Big Data therefore contains a contradiction. On the one hand, it reduces individuals to quantifiable bits of information – demographics, consumer choices, passport-ready identity
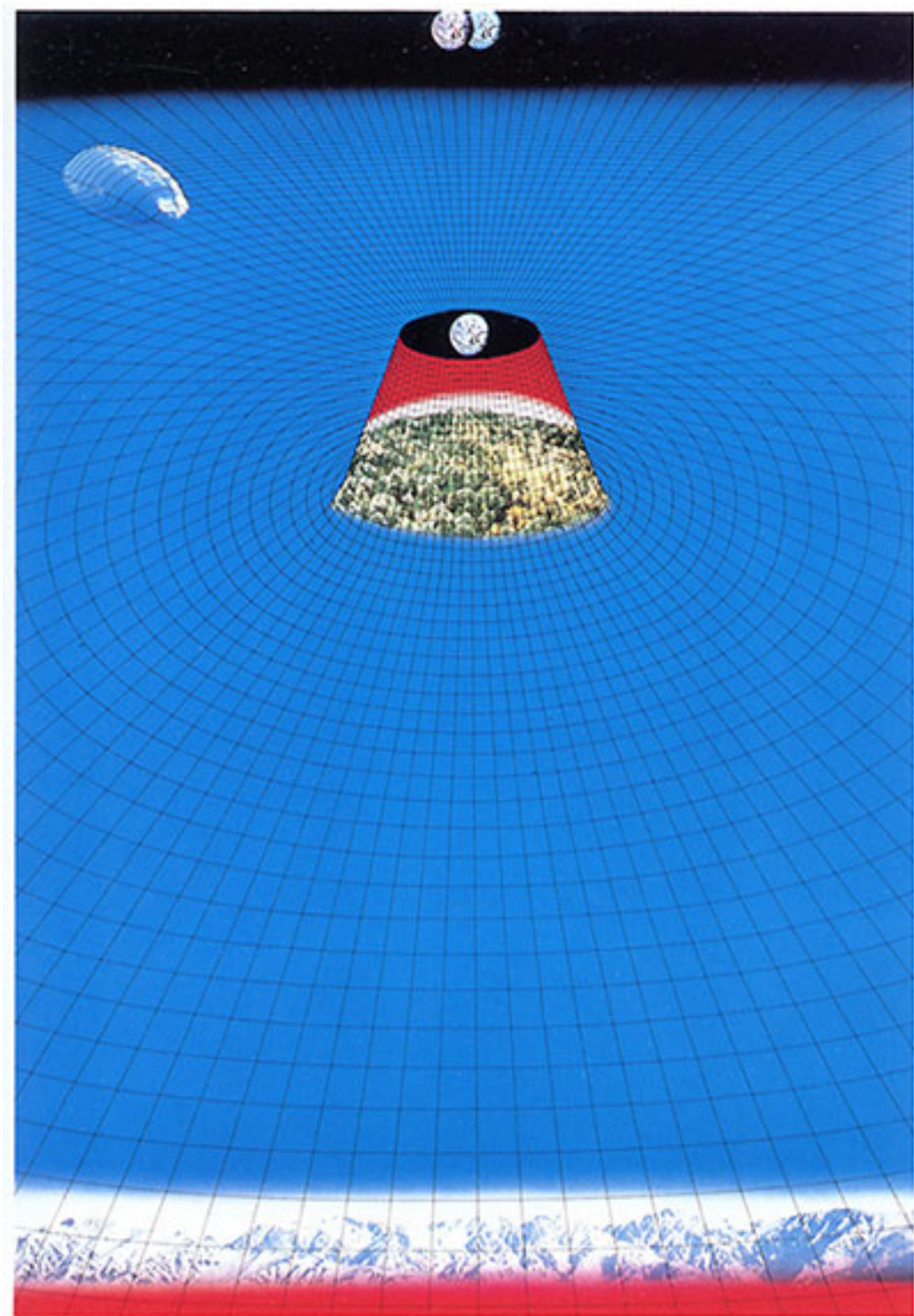
Japanese designer Kazumasa Nagai's surreal depiction for a poster design, late 1970s.

markers. On the other hand, Big Data exists as an endless stream of unprecedented scale, aggregating flows of people, their places, things, and activities into ever larger undifferentiated masses. Big Data therefore instantiates Borges's oscillation between map and territory as a permanent feature of society. It is a concrete instance of the social as such, a manifestation of the longstanding and active ambivalence in the categories, concepts, and ideas that arbitrate the relationship between individuals and the social world. This ambivalence is especially clear in the rapidly developing field of the "Digital Humanities," an uneasy hybrid of the humanities and the sciences that negotiates the relationship between map and territory, self and society, by appealing to the Janus-faced enigma of data.

A case in point is Selfiecity, an ambitious online project launched in February 2015. This site attempts to provide some kind of map for the territory staked out by the selfie, that now pervasive form of self-portraiture that has garnered an exponentially growing amount of attention since the term was deemed "word of the year" in 2013. Art historians and cultural critics have competed to offer in-depth analyses. Julian Stallabrass penned a genealogy in the *London Review of Books* in 2014 ("Most selfies are pastiche and many tip into parody."), and there was an academic conference, "Imag(in)ing

the Self in Digital Media," in Marburg, Germany in April 2015.[1]

To some, the emergence of the selfie reflects the sheer narcissism of youth; to others, it empowers individuals with the means for more self-expression. Still others see the significance of the selfie in its technological base: they argue that cellphone cameras, along with constant and easy access to Instagram and Facebook, democratize both the making and distribution of images, while encouraging complete conformity in style, peer-to-peer. Finally everyone can be an artist – or at the very least, an image- and trend-maker – so long as they adhere to a discrete stylistic repertoire.

Selfiecity is a welcome intervention into these cultural diagnoses because it telescopes out from *the* selfie to inquire after the *networks* of selfies. Motivating Selfiecity's method is a theoretical, even ethical, question: How do we define and express the position of the individual (and their agency) in relation to Big Data, which attempts to encompass all of our social interactions? The most compelling – and the most troubling – part of the project is not the selfies, but how the project appeals to data visualization to navigate this challenging question.

Selfiecity is an interdisciplinary and collaborative endeavor between Lev Manovich,

Since May 2015, Tom Bittmann, a teenager and founder of the app Wall of Selfies, holds the Guinness record for the largest selfie in the world with 2529 people portrayed in the background.

the new media theorist, historian, and director of the Software Studies Initiative at the CUNY Graduate Center, and a team of university-affiliated and independent researchers from fields as disparate as computer science and art history. The researchers created their data set by selecting selfies generated by Instagram users in five cities across the globe – Bangkok, Berlin, Moscow, New York, and São Paulo – with 640 selfies from each. (A recent installation of the project this winter added a sixth city, London, to the mix.) They chose photos from an initial sample of 120,000 randomly selected Instagram images whittled down by "Amazon Mechanical Turks." These are neither mechanical, nor, necessarily, Turks, but un-predicated humanlaborers open to completing odd and interesting tasks and who are connected with potential employers by the online retailer. They are called Mechanical Turks after the eighteenth-century chess-playing machine that, in the history of automation, has become an emblem of the synthesis of man and machine.

Instagram time- and location-stamps its images, and the Amazon laborers guessed the age and gender of the person in each selfie. Then the core research group ran the images through face analysis software, which provided algorithmically calculated estimates of head tilt and rotation, position of facial features (eye, nose, mouth), as well as marking the presence or absence of glasses. The resulting 3200 images contained all this metadata of time, place, age, gender, and formal composition.

Selfiecity subjects this sample of 3200 selfies to data visualization, statistical analysis, and historical and theoretical reflections, displaying the results in a variety of ways. The "findings" section contains bar and line graphs illustrating results – e.g., that more women than men take selfies and strike more "extreme" poses (when extremity can be measured by head rotation), that more young people take selfies (the average age is 23.7), and that people in Moscow smile less than those in Bangkok. In this section, individual selfies are synthesized, grouped according to ready-made categories of identity and nationality. Another section containing visualizations of the data called "image plots" organizes the material into stylized patterns, such as a gridded cube of all the selfies from each city organized by head tilt, a series of graphs showing "smile distribution" according to gender and city, and another set of charts showing the gender and age breakdown of selfies, also separated by city. Another section of the site offers an interactive "selfiexploratory" component, in which users themselves can sort the data by place, age, pose, mood (calm, angry, happy), and features (glasses - yes/no; eyes -

open/closed; and mouth - open/closed). Finally, the "theorizations" include essays that offer an art-historical analysis situating the selfie within the history of self-portraiture and vernacular photography, a manifesto-like meditation on "imagined data communities," and a critical assessment of the project itself from a feminist perspective that, among other things, acknowledges the rigid gender binary the project participates in and points to the problem – by no means confined to Selfiecity – that the actual data collection is understood to be so menial as to be outsourced to relatively unskilled and low-wage laborers. But even as these essays seem to tackle the network of selfies that the data illustrates, they take social significance of the form (and its analysis) as something given, something that exists already, out there in the world, sui generis, rather than acknowledging the extent to which the project itself posits that value, let alone making an argument for why we should agree. It is Selfiecity's methodology – or lack thereof – and the way it tacitly constructs both analytical value and an image of the social that demands further analysis.

In working across disciplines and developing new methods for research, Selfiecity is exemplary of the Digital Humanities. Methodological innovation is, according to Manovich, "the key question of digital humanities – how to combine 'distant reading' of patterns with 'close reading' of particular artifacts – by proposing a multi-scale reading."[2] Selfiecity's biggest achievement is its combination of formal analysis – the close study of compositional decisions – with maps that situate each selfie as one node amidst a wider field. In this regard, the project resembles Phototrails, another of Manovich's attempts to take on the world of social media–bound images. Phototrails tackles the whole gamut of photographs uploaded to Instagram, sorting them by hue, brightness, and upload time and creating image plots to "explore visual patterns." These patterns are more aesthetic than analytical – a number of image plots resemble rainbow-fringed black holes with swirling pixels ordered by color family. Other patterns are attached to specific events, like the tracking of image production around disasters like Hurricane Sandy.

Both Selfiecity and Phototrails therefore shuffle between offering concrete findings and taking a more exploratory approach that refuses to nail down conclusions in favor of trying out different data visualizations to no foreseeable end. The projects' multilevel strategies seem at first like just a large assortment of maps describing a territory, but the diversity and breadth of strategies allow the designers to

imagine they are not making a map at all. The data is synthesized or aestheticized. Whether radiating outward in image plots suggestive of the endless stream of selfies on social media platforms, or gravitating inwards to cohere into (somewhat intuitive) statistics on gender and age, the projects stop short of offering any conclusions, interpretations, or analysis. The focus is resolutely on the data, the software, and its vicissitudes. To understand this focus, consider the careful line toed already in Phototrails:

> we do not necessarily have to aggregate user generated content and digital traces for the purpose of Durkheim-like mapping of society where individual people and their particular data trajectories and media diaries become invisible … The individual and the particular do not have to be sacrificed for the sake of data aggregation, or "large scale patterns." Instead, we can perform "thick visualization" … of the data, practicing "data ethnography," and following individuals rather than only "society."[3]

At first, this passage seems to be an optimistic treatise about how a project can have it both ways: it can generate maps that illustrate patterns without "sacrificing" individuals because alongside these patterns are the metadata about them. And terms like "thick mapping" and "data ethnography" suggest that some generalizations will come out of the aggregation of these particulars. Yet this optimistic ideal – that one can synthesize all the methods to avoid the pitfalls of any single one – is just another form of map/territory confusion. A map that captures every individual in all his or her singularity would be no map at all. Motivating Selfiecity's use of multilevel methodologies, I am suggesting, is a category confusion that conflates the map with the data, and the data with the territory. Moreover, this conflation is motivated by an underlying *anxiety about* and *longing for* the social. A longing for some knowledge of large-scale patterns, or broader social trends, motivates the adoption of methods like data visualization and statistics. But the anxiety refuses to synthesize the results in any determinate, conclusive way. When Manovich and his coauthors mention "Durkheim-like mapping," they are not rejecting a method, they are avoiding a category of knowledge: the social fact.

Social facts are the values, norms, and habitual behaviors that are enacted and embodied by individuals but exist beyond them. They are therefore conceptual maps, ways of

shaping and giving meaning to the territory of social life. Social facts had to be created before the field of its study – sociology – could be developed. This was forged on two fronts: theory and methodology. French sociologist Émile Durkheim contributed to both, and his book *Suicide: A Study of Sociology,* published in 1897, is illustrative of how the two are integrally intertwined. *Suicide* stands as one of the first systematic applications of statistical analysis to the study of social phenomena, and it was revolutionary in showing that an apparently private, individual decision – to choose to live or die – correlated overwhelmingly to certain social factors. Suicide, Durkheim found, was especially prevalent in societies in which an individual felt insufficient distinction between themselves and their community – such that feelings of shame or dishonor, for example, became too much to bear. It was also prevalent when the converse obtained – when an individual felt too little connection to those around them, succumbing to loneliness and alienation. (At this early date, Durkheim had to calculate the numbers manually, making up much of his method as he worked and all the while relying on the help of his students, among them the anthropologist Marcel Mauss.) From an inchoate mass of data, replete with all the particulars and quirks of individual situations, Durkheim shaped discrete categories that all led him to one conclusion: that suicide occurs when an individual's sense of belonging in society is unbalanced. Public life lay at the heart of what had previously seemed most private.

Statistical sociology enabled Durkheim to posit the existence of social facts, which "consist of manners of acting, thinking and feeling external to the individual, [and] which are invested with a coercive power by virtue of which they exercise control over him."[4] Once isolated, these facts could support wide-reaching generalizations about society and its constituent parts. This desire to stabilize the social as an object of knowledge was a response to the rapid disintegration of traditional social institutions in the crucible of an accelerating modernity. In the face of these forces, statistical sociology offered a theory of society in the face of its apparent dissolution. While Durkheim has been (and should be) criticized for making maps that inadequately represent the complexity of the territory, this criticism is itself indebted to his singular insight, extending its logic in admitting the significance of more social facts than he could see. Durkheim is a reminder that statistics and data never stand alone.

Selfiecity, on the other hand, employs statistics to the exact opposite end for which they were intended: to supplant the social fact rather than assert it. That is, the project figures

the statistic as the end rather than the means, and in so doing disfigures whatever it is that is social in its aggregated facts. This confusion informs Selfiecity's form and content alike, evinced in the two types of "findings." On the one hand, there are statistical results, Durkheim's charts without the analysis or interpretation: bar charts of women versus men, age breakdowns, smile distribution. These are presented as if they speak for themselves, though in fact they beg for further analysis and deconstruction. On the other hand, there are the claims to indeterminacy and experimentation: aesthetically compelling image plots, texts promising knowledge yet to come but not yet realized, meaning prefigured as patterns. Each strategy tempers the other, and in both cases the design of information replaces its interpretation. Selfiecity exploits data's chimerical character, functioning more like a new media art project than a sociological study. It renders statistics an aestheticized experimental form rather than a mode of analysis. In so doing, the project points to a reason for statistical data's renewed appeal. Data can invoke a totality in a way that is not totalizing, it can create an image in which individuals and "society" remain as unsettled and fluid as the image stream of online media that so often mediates between them today. But this is a non-position that ultimately leaves us with no navigational tools, a view from an imagined outside that predictably secures a vision from nowhere.

So we should be warned: although data is neither map nor territory, it can foster their confusion. It is a seductive mode of representation that can easily trap an intellectual milieu terrified by representation, providing a method for running away from its history and its own activity in the present. What data means – how it is interpreted, and to what ends – has implications not only for privacy and security but also for how we exist and understand our position as humans in the world. What is obfuscated by Selfiecity's fetish for methods, then, is not only the social but the power that maps, both conceptual and literal, have in shaping it. As thinkers, critics, artists, and investigators, we have an obligation to methodology, and we need to remember that this includes not only attention to the means we employ, but the ends to which they work.

✕

1
Julian Stallabrass, "On Selfies," *London Review of Books,* Vol. 36, No. 11 (June 5, 2014) http://www.lrb.co.uk/v36/n11 /julian-stallabrass/on-selfi es. For another account, see Jerry Saltz, "Art at Arm's Length: A History of the Selfie," *Vulture*, January 26, 2014 http://www.vulture.com/2014/ 01/history-of-the-selfie.htm l.

2
Nadav Hochman and Lev Manovich, "Zooming into an Instagram City: Reading the Local Through Social Media." *First Monday*, Vol. 18, No. 7 (July 1, 2013) http://firstmonday.org/artic le/view/4711/3698

3
Ibid.

4
Émile Durkheim, *The Rules of Sociological Method: And Selected Texts on Sociology and its Method*, ed. Steven Lukes, trans. W. D. Halls (New York: Free Press, 2014), 21.