Autographic Design

The Matter of Data in a Self-Inscribing World

Dietmar Offenhuber

Autographic Design

metaLABprojects

The metaLABprojects series provides a platform for emerging currents of experimental scholarship, documenting key moments in the history of networked culture and promoting critical thinking about the future of institutions of learning. The volumes' eclectic, improvisatory, idea-driven style advances the proposition that design is not merely ornamental, but a means of inquiry in its own right. Accessibly priced and provocatively designed, the series invites readers to take part in reimagining print-based scholarship for the digital age. www.metalab.harvard.edu

Series Editor

Jeffrey T. Schnapp

Autographic Design

The Matter of Data in a Self-Inscribing World

Dietmar Offenhuber

The MIT Press
Cambridge, Massachusetts
London, England

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Offenhuber, Dietmar, author.
Title: Autographic design : the matter of data in a self-inscribing world / Dietmar Offenhuber.
Description: Cambridge, Massachusetts : The MIT Press, [2024] | Series: metaLABprojects | Includes bibliographical references and index.
Identifiers: LCCN 2023007789 (print) | LCCN 2023007790 (ebook) | ISBN 9780262547024 (paperback) | ISBN 9780262377089 (epub) | ISBN 9780262377072 (pdf)
Subjects: LCSH: Information visualization. | Graphology. | Evidence.
Classification: LCC QA76.9.I52 O35 2024 (print) | LCC QA76.9.I52 (ebook) | DDC 001.4/226—dc23/eng/20230317
LC record available at https://lccn.loc.gov/2023007789
LC ebook record available at https://lccn.loc.gov/2023007790

Introducing Autography

> Nature will be reported. All things are engaged in writing their history. The planet, the pebble, goes attended by its shadow. The rolling rock leaves its scratches on the mountain; the river its channel in the soil; the animal its bones in the stratum; the fern and leaf their modest epitaph in the coal. The falling drop makes its sculpture in the sand or the stone. Not a foot steps into the snow or along the ground, but prints, in characters more or less lasting, a map of its march. Every act of the man inscribes itself in the memories of his fellows and in his own manners and face. The air is full of sounds; the sky, of tokens; the round is all memoranda and signatures, and every object covered over with hints which speak to the intelligent.
>
> —Ralph Waldo Emerson, *Goethe*; or *the Writer*[1]

The Grossglockner, the highest mountain in Austria, is conveniently accessible through an alpine road. From there, a small funicular train leads down to tongue of the largest glacier in the eastern Alps called the Pasterze. In 1960, the funicular delivered riders right onto the glacial ice. When I stepped out of the train car in the summer of 2022, I found myself in the middle of a mountain slope with a lake far below me, glacial ice only barely visible in the far distance. To my right was a sign saying "Gletscherstand/Glacier Position 1960." As I followed the path down, I encountered many similar signs marking the surface of the glacier in 1965, 1970, 1985, and 1990. At the 2005 mark, I still hadn't reached the lake, and the ice was still far away (figure 0.1). Every year, the trail grows in length by about 20 to 50 meters.

It is difficult to imagine a more vivid visualization of a warming environment, comprising the retreating glacier, the physical effort of hiking, and the disappointing tease of the funicular that now covers only a third of the way. The historical signposts play a crucial role in this visualization. While the glacier shows what is left, the signs show what is already gone—they help me

Figure 0.1
Pasterze Glacier, Austria. For scale, note the group of hikers on the shore to the right of the signpost. Photo by the author.

imagine the shape and boundary of the ice volume that has disappeared. But beyond evoking a mental image, the signs enable me to see concrete features of my surroundings differently: the reason why the funicular ends halfway, the subtle differences in the landscape that allow me to guess where the ice once ended, like the former shoreline of a half-empty reservoir etched into the terrain. The historical data points frame my experience and guide my attention. And yet I wonder how my experience might have been different if someone had secretly moved the signposts or changed their dates. The data are inseparably tied to the glacier, while the experience of the latter is partially shaped by the numbers displayed on the signs.

But besides serving as a visceral display of a heating climate, the Pasterze trail offers another lesson. Much of what we know about the climate, we know through computational models operating on vast amounts of data. They allow scientists to study phenomena that cannot be directly observed. Beyond these data inscriptions, however, the world itself is self-inscribing, or *autographic*. The physical environment is not a passive canvas for digital data layers; it is actively recording and processing information. As Ralph Waldo

Emerson wrote in 1850, "All things are engaged in writing their history"—the environment bears the marks of past events. But in contrast to digital data sets, these inscriptions do not take the form of a description—information is embodied in the phenomena themselves, often in opaque and illegible ways.

As the relationship between the signposts and the glacier indicates, data and the lived experience are entangled in ways that are not always obvious. This book is for anyone interested in data, visualization, and design—an audience for whom digital information has largely become second nature. Many information designers and data scientists are motivated by helping their audiences make sense of today's biggest challenges and navigate a post-truth media environment. However, to achieve this, I believe that it is sometimes necessary to turn away from digital models and engage the material world with all our senses. This book presents autographic design as a countermodel to data visualization—a practice that is less concerned with interpreting data than with revealing their material origins and the relationship between data and the world.

Symbolic Information and Its Discontents

Digital computation involves encoding the world into symbols.[2] Practically speaking, this means associating people with characters such as "m" or "f," or describing the atmosphere using variables such as temperature, humidity, and atmospheric pressure. None of these categories can capture the nuances of what they aim to represent, and they frequently prove to be inadequate. And yet computation benefits from such simplifications. They focus observations and give them consistency and direction.

Symbols can be translated and combined in endless variations, regardless of whether they stand for a person, a program instruction, or a calculation result. Such transformations never add new information,[3] and yet they have produced a complex digital world whose inner workings are increasingly opaque. Outside information comes from sensors and other forms of data capture, often compared to a natural resource: extracted from the environment, harvested from users, distilled into different formats, fed into machinery, and reclaimed even from its exhausts. But despite these corporeal metaphors, data are generally presented as abstract, virtual, and unaffected by the frictions of the physical world.

Data are commonly experienced through charts and graphs, which have become the native media for observing the economy, the heating of the climate, or the opinions of voters. If data are considered symbolic representations, visualizations are representations of representations. Categories can be mapped to a set of colors, and quantities encoded into bars of varying lengths. According to the canon of data visualization, the visual vocabulary should follow the data type—such as numbers or categories—not what the data represent.[4] This separation of form and content again reinforces the idea of data as abstract entities: *regardless of the topic, the visual language remains the same*.

But data visualization cannot begin unless data exist. Once they exist, however, the difficulties involved in their generation tend to be quickly forgotten. Data are not simply collected; they are actively created: data are *taken*, not *given*.[5] Nevertheless, metadata and methods sections are frequently skipped by practitioners, and many data sets are unspecific in this regard. Geographic coordinates, for instance, might have been acquired from satellites or cell tower infrastructures, or they might be inferred from street addresses. When data change hands and are compiled into other data sets, these various origin stories can lead to unforeseen consequences. A farm in Kansas received a flood of unpaid parking tickets, visits from law enforcement agents, tax collectors, and ambulances, just because it happened to be in the exact center of the country. A commercial geolocation service that infers geographical coordinates from Internet Protocol (IP) addresses has chosen this location as a default value when no other information was available.[6] When the origins of data are ignored, gaps routinely open between the data and the material world. The methods of data visualization, however, are rarely concerned with the data generation process, only with their representation.

But data are not abstract representations; they invariably manifest in physical form. To paraphrase Friedrich Kittler, there is no such thing as software, only voltage differences:[7] all code written in high-level programming language boils down to elementary electrical impulses. Visualization designers continuously struggle with the material limitations of their medium, whether it is the abysmal print quality of a scientific journal or the small size of a smartphone screen. There are good reasons to prefer the clean world of symbolic representation: in symbolic space, information does not age or degrade, and every datum can be treated equally by algorithms and statistical models. The material side of data tends to complicate things: each datum becomes a unique case, with its own history and idiosyncrasies.[8] In the material world, no two things are exactly alike.[9] Each camera sensor is unique, and the specific noise patterns in each image point to the exact camera with which it was taken. Data materialities encompass the infrastructures, bureaucracies, and labor involved in data collection, and the maintenance necessary to keep artificial intelligence (AI) systems running, often subject to exploitative labor arrangements.[10] Further upstream, we find the raw energy and geological resources consumed in the manufacturing of technologies, lithium, cobalt, and lanthanides. Downstream, we find the toxic legacies of e-waste.[11]

"The principal disadvantage of symbols is that we confuse them with reality," as writer Alan Watts observes.[12] One can easily get lost in the world of digital media, making it difficult to imagine information that is not symbolically encoded in words, pixels, or bytes. And yet the material origins and manifestations of data matter. Without considering them, a fabricated data set may be indistinguishable from an authentic one. As public discourse is full of controversies around facts and counterfacts, the gaps between data and the world become increasingly troubling. These controversies often revolve around how data were generated and facts were established—something that traditional visualization methods don't account for. Even in less contested situations, reconciling a statistical model with personal experience is difficult. This is where some of the greatest challenges lie.

Global warming has often been described as invisible because it is too slow or too big for direct experience. As each new year brings record-setting temperatures, floods, and forest fires, this argument no longer hits the mark. Nonetheless, the climate is not a physical phenomenon but a statistical concept—it is the history of the weather based on long-term averages.[13] While one can experience the weather but not the climate, global warming reveals itself in many physical clues—from shrinking glaciers and disappearing species to the widths of tree rings, which reflect changing seasonal patterns. Beyond an ever-expanding digital universe of data, the world continuously inscribes itself in countless ways. Can we consider these inscriptions as a form of data? This book is an invitation to leave the walled gardens of symbolic information and explore its connections and relationships to the world.

Above All Else, Show the Data?

In 2010, a simple visualization started circulating among climate skeptics (figure 0.2). Its creator, the geologist and petroleum engineer David Lappi,

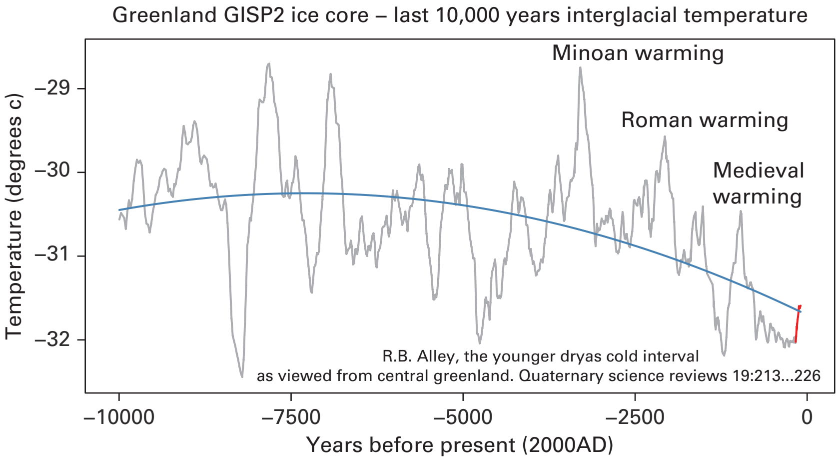

Figure 0.2
Re-creation of "Greenland GISP2 Ice Core—Last 10,000 Years Interglacial Temperature," as created and annotated by David Lappi. First publicized version from February 2010. In 2016, the erroneous date 2000 AD was corrected to 1950.

wanted to show not only that the global climate has experienced many warmer periods in the past, but also that its most significant threat at the moment is global cooling rather than warming. The chart has been reproduced in many variations[14] and most recently was prominently featured in a publication by the conservative Heartland Institute, which distributed it to virtually every science teacher in Florida.[15]

In Lappi's original chart, a recent rise in temperature is visible, but it is dwarfed by warmer periods during the Middle Ages and antiquity. The visualization is based on a publicly available paleoclimatological data set published by a team led by the prominent climate scientist Richard Alley, which contains temperature estimates for Greenland covering the past 50,000 years.[16] It is worth examining the original data set and its patterns; in doing so, I will follow Edward Tufte's guiding principle of information design: "above all else show the data."[17] The public data set consists of a plain text file following a descriptive header, starting with two data columns labeled as "Age" (in thousand years before the present) and "Temperature" (in degrees Celsius):

DATA:

1. Temperature in central Greenland

Column 1: Age (thousand years before the present)
Column 2: Temperature in central Greenland (degrees C)

Age	Temperature (C)
0.0951409	–31.5913
0.10713	–31.622
0.113149	–31.6026
0.119205	–31.6002
0.119205	–31.598
0.125451	–31.6656
0.132407	–31.7235

. . .

Using software such as Microsoft Excel or the open-source statistical software R, one can easily produce a line chart from the data set; it will look similar to the one shown in figure 0.3. We see that Lappi used only a small part of the data set. Its larger shape reflects the dramatic temperature swings in Greenland during the last Ice Age, which ended about 12,000 years ago.

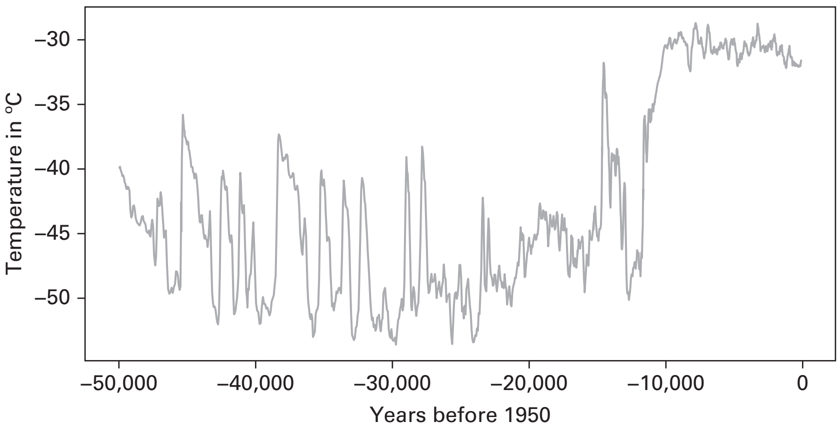

Figure 0.3
Plotted temperature estimates from Richard Alley's GISP2 Ice Core data set. *Source*: Alley 2000.

Zooming in on the last 10,000 years of the data yields a chart that is almost identical to Lappi's visualization, which included a few additional annotations and a trend line. As the trend line indicates, temperatures are decreasing and seem to accelerate while doing so: "the long-term temperature trend on the planet appears to be down, not up," as Lappi notes.[18]

But the most conspicuous detail is the emphasis on the last section of the curve, indicating a temperature increase at the end of the data set, eclipsed by earlier warming periods. This last section of the curve resembles what became known as the "hockey stick graph" included in the 2001 Intergovernmental Panel on Climate Change (IPCC) report and further popularized in Al Gore's film *An Inconvenient Truth.*[19] Lappi suggests that the two charts show the same information, but while the original hockey stick graph covers only 1,000 years, Alley's data set (published a year later) provides a larger picture, which reveals that the warming is much less dramatic.

It should come as no surprise that climate scientists have many problems with this interpretation, as Lappi equates the local temperature of Greenland with the global climate. But he also undermines his argument with a data error that is not immediately obvious. His last datum refers to a time ninety-five years before the present, which Lappi implies to be the year 2000, the year of Alley's publication. This would mean that the data set would end in the year 1905, but unfortunately, this is not correct either. The *before present (BP) timescale,* a convention mainly used in archaeology and geology, defines the present as the year 1950, moving the most recent datum to the year 1855. With most of the Industrial Age not covered, the data cannot be compared to the hockey stick chart, which includes data up to the year 2000.

Lappi died in 2011, but his chart continues to circulate on the Internet. Although the timescale has since been corrected, the chart is still cited by climate change deniers as evidence of a highly volatile climate beyond human influence. With the dating issue out of the way, what is troubling from a data visualization perspective is that the chart contains no major "data crimes." The data set is from a reputable source, correctly plotted, labeled, and cited—it follows Tufte's principle of "above all else show the data." The curve shows a compelling pattern of falling temperatures, but it omits the dramatic changes that have occurred over the past 150 years. The geological timescale visually diminishes the significance of recent developments. It is likely that a naive viewer, focusing only on the patterns in the data set, might notice the same cooling trend, the anticlimactic end of the chart, and reach the same conclusion as Lappi. The chart requires context and an understanding of the dating conventions in climate science and the origins and limitations of the data sources. Data cannot be neatly separated from their context.

Information designers will point out that complicated matters require more nuanced explanations. Sometimes we need more elaborate visualizations that highlight the relevant aspects and add annotations to clarify their relevance. I fully agree with them—Tufte's maxim does not mean "*only* show the data." But such explanations require going beyond the data content and examining how it relates to phenomena and the people and mechanisms that generated them. Following such a route will inevitably force us to deal with the material nature of data and the conditions within which they are generated.

Traces and Latent Information

We can return to Richard Alley's data set, a text file of a modest 116 kilobytes, to explore the material substrate of the data. It is based on Greenland ice core records, which, as its header notes, "provide an exceptionally clear picture of many aspects of abrupt climate changes."[20] Past climatic conditions are inscribed as traces into the ancient ice in many ways, rendering it a useful proxy—the next-best data source for climate scientists who cannot directly measure temperatures of the distant past. The Greenland Ice Sheet Project 2 (GISP2), with Alley as the principal investigator, extracted a 3-kilometer-long ice core taken from the highest point of Greenland's glacier, from where the researchers drilled down into bedrock. It is currently stored at the National Ice Core Laboratory in Lakewood, Colorado. Alley discusses a variety of methods available to reconstruct past climates from the recovered ice in detail.[21] He notes the well-preserved annual layers—like tree rings—resulting from seasonal variations of snowfall (figure 0.4). They can be visually counted to date sections of the core and infer the amounts of precipitation.[22] Dark layers of dust point to volcanic eruptions and serve as convenient references for aligning the core with other data sources. Alley developed a visual technique for exposing seasonal layers in snow pits—two adjacent holes on the glacial surface separated by a thin wall of snow. After covering one pit with a makeshift roof, sunlight filters through the translucent wall between the pits and reveals its layers. Alley notes that lay visitors in the snow pit were able to see the layers without instruction and were deeply moved by the beauty of the blue light filtering through the snow.[23]

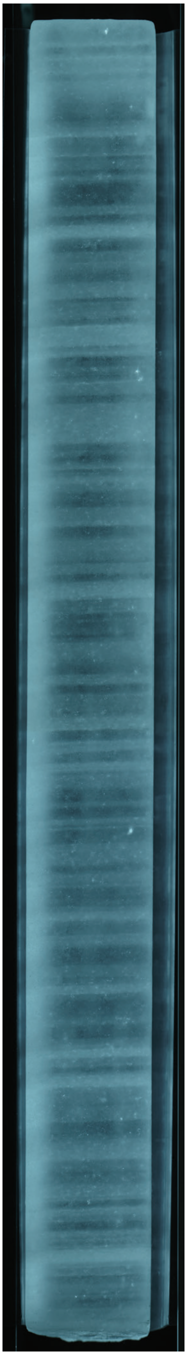

Figure 0.4

The GISP2 ice core section showing annual layer structure, back-illuminated. The core is from a depth of 1,840 to 1,841 meters. The approximate age at that depth is around 16,000 years before present (with the present set as 1950). Photo courtesy of the National Science Foundation Ice Core Facility (NSF-ICF).

Other physical qualities of the ice are less conspicuous. Aerosols and dust particles point to prevailing wind conditions, while trapped gas bubbles indicate the atmospheric composition, including levels of carbon dioxide (CO_2).[24] The reconstruction of past temperatures takes multiple phenomena into account. Alley describes a surprisingly direct approach: lowering a thermometer into the borehole reveals residues of the cold from the last Ice Age.[25] The metadata of the GISP2 data set mentions a second paleothermometer: the isotopic ratios of ice. This method takes advantage of the observation that seawater contains a stable amount of "heavy water"—molecules composed of heavier oxygen isotopes, which can be detected by a mass spectrometer. Since those molecules are the last to evaporate from the sea and the first to condense into rain and snow, the amount of heavy water in precipitation can serve as a temperature indicator. According to Alley, the proxy methods complement each other—isotopic ratios capture more detail but need to be calibrated with the less detailed but more stable data from the borehole.[26]

This brief exposition does not come close to listing the catalog of techniques for analyzing ice cores. What is striking is that a surprisingly detailed picture of the past can be extracted from a few centimeters of frozen water. While most methods require complex instrumentation and mathematical models, some seem elementary, sensory, and experiential, such as counting layers in snow pits. Many proxies offer tangible and vivid clues about the effects of climate change. Ice cores are joined by tree rings, coral skeletons, geological sediment, fossilized pollen, and insect remains, constituting a veritable material archive of traces.[27]

If climate is the history of weather, these traces are the materialized records of said history; their accumulations are analogous to statistical averages. Often, they are also veritable visualizations. The segmented and backlit ice cores are vivid displays that allow an appreciation of the nature, richness, and limitations of the data set.[28] Like data visualizations, the traces are shaped by human decisions, machinery, and scientific models. But unlike visualizations, they are not representations with assigned semantic meanings. As I will argue, cultivating the sensory experience of material data can enable a productive mode of critical inquiry into the process of data generation.

Material Data

As I will examine further in chapter 1, climate proxies are more than material data sources; we can think of them instead as *data in physical form*. Why does this distinction matter? Traditionally, data are understood as recorded observations by humans or machines. Such records do not have to be textual; archaeologists, for example, may use the term "data" to refer to artifacts extracted from a site. Critical data scholars often focus on the epistemological issues surrounding data: how agendas and worldviews inform data generation and their use as rhetorical instruments of power with associated claims of objectivity and truth. From an epistemological perspective, data are products of an interpreting human mind. Yet there is also an ontological perspective on data that is less concerned with what a datum means or communicates and more interested in how it acts.

From an ontological perspective, data can be seen as patterns and differences in the world, independent of human interpretation. The philosopher Luciano Floridi describes a datum as "lack of uniformity" in a given context,[29] echoing the cyberneticist Gregory Bateson, for whom a bit of information is the "difference that makes a difference."[30] Geologists have regarded the strata of the Earth as a natural archive that preserves the temporal and spatial relationships of its history.[31] Some physicists even argue that information itself should be considered a physical entity, with quantum mechanics imposing hard limits on what can be theoretically computed.[32] The patterns of the ice core have no semantic meaning per se, but as the causal effects of atmospheric movements, they are meaningful in a different sense.

Characterizing data as human-independent patterns in the world, and therefore objective, means asking for trouble. Framing data as entirely human independent certainly does not seem appropriate. After all, ice cores are the product of human-made machinery, underpinned by scientific models and shaped by human decisions. I want to use the term "objective" not as a shorthand for observations from a neutral position that offer access to indisputable truth, but in the weak sense of being object-like. As objects, ice cores embody the scientists' questions, as well as the technical knowledge of the drillers. They are the products of a sociotechnical project that has been developed over decades in multiple fields of research. The layers of snow glowing in a beautiful blue light can reveal themselves only by digging a snow pit. Before tree rings can present themselves, one has to cut into a tree.

Shifting the concept of data from the abstract-symbolic into the material world involves changing our focus from a long tradition of interpreting symbols to the elementary modes of manipulating and transforming matter, tuning into phenomena.[33] Meaning is not a prerogative of human cognition alone: consider how many animals are far superior to humans in detecting traces and how plants can sense changes in their environments. The meaning of traces is situated and relational, residing in the interrelationships between ice layers, dust particles, and the equipment used to extract them. As the glacial ice sample is vaporized in the mass spectrometer to determine isotopic ratios, Geoffrey Bowker's metaphor of data that are never "raw" but should be "cooked with care" becomes literal.[34]

Introducing Autographic Design

Many disciplines, including the humanities, social sciences, and design, have recently witnessed a material turn, aiming to supplant the reliance on textual metaphors and mind-body dualisms so predominant in Western thinking. The material turn also has left its marks in the area of human-computer-interaction (HCI), where embodiment, relationality, and materiality have become focuses of attention.[35] The data visualization discipline, in contrast, is still largely based on theoretical models that privilege a purely representational perspective that treats data and other information as abstract, universal, and context independent.

This is starting to change. Over the past ten years, many designers and visualization researchers have started exploring physical expressions of data through tangible objects, embedded displays, and even food.[36] Personally authored data sets and hand-crafted visualizations, exemplified in the Dear Data project, a postcard exchange in the form of hand-drawn data visualizations, have gained in popularity and challenged visual conventions.[37] Feminist critiques of data visualization question its implicit claims of authority and objectivity and emphasize the context and situatedness of data.[38]

Among these developments, the recently formed field of *data physicalization* studies the advantages of expressing data in three-dimensional, physical form over conventional screen-based visualizations. The researchers Yvonne Jansen and Pierre Dragicevic frame data physicalization as an extension of the data visualization palette, saying: "Traditional visualizations map data to pixels or ink, whereas physical visualizations map data to physical form."[39] In their research, they found indications that people are better at interpreting data when they are able to touch them rather than just look at them. But Jansen and Dragicevic's interest goes beyond questions of performance; it also includes the rich cultural history of physical visualization, from the French cartographer Jacques Bertin's physical matrix displays to the role of physical models and prototypes in the history of science and engineering.[40]

Design practices that transform data into objects open new ways to explore data sets, and yet they do not address the previously described blind spot of data visualization: the hidden material circumstances of data generation. To address this issue, I propose *autographic design* as a countermodel to data visualization. It describes a set of techniques for transforming phenomena into material traces and guiding their interpretation.[41] The term "autographic" (meaning "self-writing" or "self-inscribing") refers here to the self-inscribing qualities of ice cores and tree rings, in which climate history presents itself. Autographic design is less concerned with the different ways in which data can be visualized and more with how the physical qualities of an entity (e.g., an ice core sample) can be revealed and transformed into a data set. In contrast to data physicalization, it does not translate a data set into physical form but focuses on the physical forms at the basis of data generation. Autographic design is not limited to visual phenomena; it includes other sensory modes such as hearing, feeling, and smelling. The term "countermodel" is not meant to set autographic design in binary opposition to data visualization. There are considerable similarities and overlaps between the two domains. The comparison is meant to clarify the characteristics, assumptions, and limitations of both models.

At the heart of autographic design are material traces and the practices involved in seeking, revealing, preserving, transforming, and presenting them as material evidence. The goal is often, but not always, to transform them into symbolic data records. But autographic design is also interested in the opposite: examining artifacts and glitches in data sets as unintentional traces of data generation.

Designing autographic visualizations means shaping the material conditions that will allow a trace to present itself and provide a framework for its interpretation. The design process can be imagined as a process of tuning into a phenomenon, similar to how the body of a violin is adjusted to capture the resonant frequencies of its strings. Autographic design is hence better understood as a practice rather than a thing. John Durham Peters describes the sundial as the prototypical autographic device—a self-inscribing mechanism that foreshadowed the wave of autographic instruments created during the eighteenth and nineteenth centuries, ranging from weather clocks to the photographic camera.[42] The use of the terms “autographic” and “self-registering” in patent applications peaked, as a Google n-gram search indicates, at the turn of the twentieth century.[43]

By “trace” in this project, I mean things that are typically described as symptoms, markers, signs, signatures, signals, traces, clues, indices, and indicators. I use the compound term “trace-phenomena” because I want to capture ephemeral impressions as well, such as smells or sounds. This broader meaning has two rationales: first, the differences in terms of permanence between a footprint in fresh concrete, the warm spot on a recently occupied chair, or the disturbances of air perceived as sound are only gradual; and second, traces are not monolithic entities but rather assemblages of many components, some more ephemeral than others. Even the familiar animal track encompasses a wide range of interrelated clues. As the South African anthropologist Louis Liebenberg observes, an experienced tracker takes many clues into account and considers a track (*Spoor* in Afrikaans) as undergoing constant change:

> In the narrowest sense of the word “spoor” simply means “footprint,” but in tracking it has a much wider meaning, including all signs found on the ground or indicated by disturbed vegetation. Tracking also involves signs such as scent, urine and feces, saliva, pellets, feeding signs, vocal and other auditory signs, visual signs, incidental signs, circumstantial signs, blood spoor, skeletal signs, paths, homes and shelters. Spoor are not confined to living creatures. Leaves and twigs rolling in the wind, long grass sweeping the ground or dislodged stones rolling down a steep slope leave their distinctive spoor. Markings left by implements, weapons or objects may indicate the activities of the persons who used them, and vehicles also leave tracks.[44]

Autographic design is concerned not only with finding traces but also with techniques of trace-making, conceiving mechanisms to record ephemeral phenomena as permanent traces. While scientific practice has long relied on skills and techniques for turning invisible phenomena into visible traces, many of these techniques have since been replaced by digital methods. What, then, is their relevance in the early twenty-first century? As we will see, many recent developments indicate a renewed interest in autographic phenomena.

New Analog Practices

Data visualizations have become the default for dealing with almost any matter of concern. And yet we find a growing number of citizen scientists, environmental activists, and forensic amateurs who use decidedly analog methods for investigating and constructing evidence of pollution, global warming, and disinformation. Often without institutional affiliations that inspire trust in the data they produce, they may use material displays to demonstrate causality and present evidence. The same strategy can help engage participants and collaborators by staging data collection as a public experiment.[45] Autographic displays do not present conclusions but invite curiosity and causal reasoning, engaging the audience differently.

In digital forensics, the material qualities of data play a crucial role in securing digital evidence and uncovering disinformation—from identifying fake online identities based on profile pictures that bear the characteristic artifacts of AI image generators to the digital traces left on hard drives and in the corners of the dark web. To detect manipulations of data, digital forensics experts have developed methods that focus on the unintentional traces generated during the production of a data set rather than its explicit content.[46]

In computing, the limitations of the digital paradigm have been emerging over the past decade. Chip design is increasingly approaching its physical limitations, which is why interest has revived in analog computers and in predigital methods of modeling and visualization.[47] Computer scientists studying the behavior of artificial neural networks—convoluted structures that resist a top-down analysis of their inner workings—are currently rediscovering classic, empiricist approaches that probe the network's behavior in ways not unlike those of a historical naturalist probing a natural phenomenon.[48]

Autographic properties and material agency also have become focal interests in material science and microbiology. Material researchers work on smart materials with autographic qualities that are intentionally designed. Such materials may have the capacity to respond to environmental conditions, change their appearance, or self-assemble into different configurations. Biodesigners have similar goals, modifying bacteria DNA so that they can indicate the presence of toxins and other substances.

Comparing the Models

Table 0.1 briefly summarizes the main differences between data visualization and autographic design:

Table 0.1
Comparison of data visualization and autographic visualization

	Data Visualization (*Infographics*)	Autographic Design (*Autographics*)
Relationship to symbolic data	Begins with data	Ends with data
Relationship to the world	Representational: visual marks standing for something	Presentational and relational: traces presenting themselves, resulting from causal interactions
Relationship to the mind	Epistemological emphasis: visualization as a product of knowledge	Ontological emphasis: visualization as part of a physical phenomenon
Scope of inquiry	Inward: discovering patterns in a data set	Outward: revealing the process of data generation
Role of design	Mapping data to visual variables with unambiguous meanings	Elucidating the polysemic qualities of a phenomenon

Data visualization relies on a data set, which is then filtered, transformed, and mapped to a visual language. Autographic design starts with the phenomenon, which is probed and transformed so that its traces can be interpreted. These traces can then either become the basis for new data sets or serve as visualizations in their own right.

The symbols used in data visualization stand for data values, which themselves typically represent observations in encoded form. Autographic design, on the other hand, is nonrepresentational. Traces present themselves; they do not stand for anything. This distinction, however, is not clear-cut. Data visualization includes nonrepresentational aspects: some kinds of data do not stand for observations; they are by-products of a computational process. Conversely, autographic displays also can take on representational aspects. The smoke trails in a wind tunnel present themselves as traces, but they also stand for the invisible movements of the surrounding air.

The third difference focuses on how visualization relates to the human mind. In many ways, data sets are shaped by interpretations and involve abstractions and simplifications based on worldviews. Data visualizations are often considered cognitive tools, emphasizing their epistemic role. Eventually, traces need to be interpreted as well, but as objects, they are often protoepistemic: they exist before interpretation, not defined by language, rubrics, or categories. Autographic design emphasizes ontological questions such as what constitutes a trace and how its components relate to each other. Preserving and revealing traces, however, often requires instruments that embody knowledge and worldviews. For example, when facial recognition systems systematically have problems handling dark skin, the causes cannot be dismissed as unintentional material side effects. They reflect the system's design, which in turn indicates the preponderance of white engineers, their priorities, and their assumptions about their users.[49]

A fourth difference concerns the scope and direction of the inquiry. Methods of data visualization may seek patterns and correlations inside the data set in order to demonstrate something about a phenomenon that is considered "outside." Autographic design is about looking outward. It focuses on the context of data generation and the relationships among clues to learn about the origin of a data set. In so doing, it undermines the ontological boundary between the inside and the outside of data.

The main practical difference concerns the design process. A data visualization designer selects variables from a data set and maps them to a set of visual variables, including symbols, layouts, and scales.[50] An autographic designer, on the other hand, interacts with materials through design operations such as *framing*, *constraining*, *aggregating*, and *coupling*, which will be described in the following chapters of this book.

Data as Autographic Entities

Despite these differences, data visualization and autographic design are cousins with a shared history. Under the umbrella of the *graphic method*, the design of statistical charts and the production of physical inscriptions were once considered part of the same discipline. The late-nineteenth-century French physiologist and inventor Étienne-Jules Marey became known for both graphical techniques and technologies for capturing physical phenomena, including devices to visualize the human pulse, the gait of animals, the flight of insects, and the movement of air around objects. He was influenced by such earlier pioneers in statistical graphics as William Playfair, who canonized the visual language of charts in works such as the *Commercial and Political Atlas* (1805); and Charles Minard, known for cartographic-quantitative representations such as the map depicting Napoleon's 1812

Russian campaign (1869). Marey's ambitions, however, went beyond creating data graphics (which he regarded as a mere clerical exercise) and toward developing a visual approach to science, following an old ambition in the natural sciences to free research from the ambiguities of language and the limitations of human perception.[51] In his striking visualizations, Marey saw manifestations of "the language of phenomena themselves," superior to all other forms of expression. The historians of science Lorraine Daston and Peter Galison describe this ideal of automatic image-making without human intervention as "mechanical objectivity."[52]

Considering the extraordinary range of visualization methods employed in his laboratory, it is notable that Marey's visualizations share a consistent visual language that makes them recognizable. Statistical data graphics, the charts created by self-registering instruments, and the images produced by his chronophotographic motion studies[53] all present themselves as dematerialized trajectories inscribed against a dark, rasterized background, blurring the boundaries between graphics, charts, and photographic images.

In the world of analog visualization, each graph is also a physical trace. A mechanical seismograph inscribes the movements of the ground onto a rotating chart drum, producing a traditional line chart. But doesn't this parallelism also apply to digital data? A datum automatically recorded by a sensor and written onto a hard drive is a physical imprint. It is equivalent to the trace created by the seismograph not only because it is a physical inscription but also because it originated from a causal chain of material transformations, leading us back to the phenomenon. From this point of view, the distinction between symbolic and material data becomes irrelevant.

To add nuance to this distinction, I will turn to another definition of autography—namely, Nelson Goodman's comparison of the autographic and the allographic arts. Goodman proposed: "Let us speak of a work of art as autographic if and only if the distinction between original and forgery of it is significant; or better, if and only if even the most exact duplication of it does not thereby count as genuine."[54] Allographic arts, on the other hand, are based on notational systems that do not rely on a single unique original; Goodman includes poems, literature, and sheet music in this category. However, he complicates this seemingly simple distinction with examples of autographic acts in systems of mechanical reproduction, such as etching the plate in the process of printmaking. Goodman concludes that the distinction between allographic and autographic arts is not a question of whether they result in a singular, unique object or multiple copies. Instead, the integrity of an allographic artwork depends on what he describes as the "sameness of spelling," the exact correspondence in symbolic space.[55]

Applied to data, a digital data set is allographic in the extreme: it can readily be transformed until it no longer shares any resemblance with its initial format, but it does retain the same meaning or value. Variables can be renamed, and values can be reformatted and normalized—census data, for instance, have to undergo many such transformations until they become useful for analysis.[56] What, then, is an autographic data set? As Matthew Kirschenbaum argues, the allographic nature of digital media amounts to a thin veneer of error correction mechanisms that overlays the deeper, "autographic but illegible signatures of actual computational inscription."[57] But it is not just the storage medium that relies on autographic processes for its sorting and classification; the modes and circumstances of data collection also leave marks in a digital data set that can be considered autographic.

To illustrate this point, let us consider another large public data set popular in big data challenges: the records of hundreds of millions of taxi trips in New York City.[58] Each trip record includes the geographic coordinates of the pickup and drop-off locations, among other data. Readers with a powerful computer can download parts of the data set and import them into a geographic information system (GIS).[59] Plotting the coordinates of the pickups and drop-offs in different colors and adjusting their transparency yields an image like figure 0.5.

Somewhat unsurprisingly, we see a map of the urban street grid emerge. Some parts appear brighter than others, indicating popular areas. But I want to call attention to other parts of the image. Some areas appear very blurry, such as the Midtown section of Manhattan. To understand this effect, one has to remember how global positioning system (GPS)[60] localization depends on the reception of satellite signals that are weak enough for a tall building to block their reception. The blurry areas are a result of such localization errors, which present as a haze when aggregated over millions of records. So, while this map visualizes only two values—longitude and latitude—it also reveals information about the three-dimensional shape of the city. Conveying this information was certainly not intentional; it is an artifact deposited by the physical limitations of GPS technology. The blurriness of the map turns out to be a workable proxy for estimating building height or the narrowness of the street canyon. To take advantage of it requires understanding the physical properties of GPS, while a more conventional approach would exclude

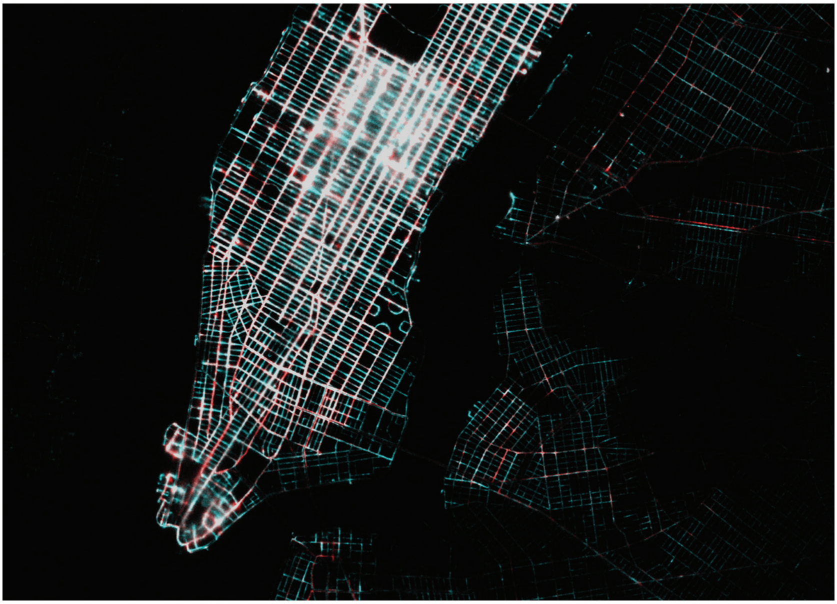

Figure 0.5
Taxicab drop-off and pickup locations in New York City. Plot by the author based on public trip record data collected by the New York City Taxi and Limousine Commission.

these artifacts as outliers. The blurry regions, like the etchings of a copperplate, are autographic signatures of the data collection apparatus. One can take the chain of self-inscription even further and think about how tall buildings are possible only in some parts of Manhattan, where bedrock is close to the surface, while other neighborhoods are built on glacial deposits that are more unstable. An echo of Manhattan's geological history is inscribed in the reported GPS locations of taxicabs.[61]

An autographic perspective on digital data does not limit itself to designated meanings; it considers all artifacts and imperfections as sources of information, even if such information often remains opaque. Any apparatus captures not only what its designers had in mind but also countless other processes and phenomena. Forensic approaches pay special attention to these unintended by-products, which offer a wealth of information not intentionally encoded.

A more general way to express this difference is by distinguishing between a representational and a relational perspective on data.[62] Based on the former, a geographical coordinate is a sign that points to a location on the planet, governed, for example, by the WGS84 standard for geographical data.[63] Nothing in this designation hints at other phenomena, such as building height. A relational perspective considers the context in which the datum was generated (its countless material relations, as well as its intended purpose). The geographical coordinate is just one intermediary data point in a long chain of material transformations. The datum is not a sign, but a signal derived from a reshaped radio wave captured from a group of satellites. The representational perspective allows reducing a data set to its values without worrying about its source, but the data set will always bear the determining characteristics of the settings and infrastructures it was generated from.[64] For this reason, Yanni Loukissas recommends that we think in terms of data settings instead of data sets.[65]

Methods

In this book, several of my previous efforts and interests converge. It continues ideas that emerged while writing my last book on waste forensics and the realization that garbage is an information resource.[66] It also builds on my work on physical and contextual information displays, often with my collaborator Orkan Telhan, in which we looked at material information from a semiotic perspective.[67] The empirical parts of the book draw from qualitative research of evidence construction in citizen science, environmental justice, amateur forensics, and digital cartography. Besides interviews and documentary research, I also examined a corpus of more than 700 examples of autographic phenomena, which served as a basis for comparative analysis. My conceptualization of traces and autographic design operations draws from literature in the areas of science and technology studies (STS), history of science, media studies, and professional literature about various disciplines concerned with traces and trace-making. Art and design projects are a vital source of inspiration and objects of investigation; I take the liberty of discussing their autographic qualities independent of whether these aspects were important to the artists who created them. I apologize for any distortions of artistic intentions that result from this cross-reading in the service of this project. Finally, to test autographic approaches in practice, I followed a research-through-design approach and developed several autographic visualizations for art exhibitions and public installations, which are also discussed in this book.

Overview of the Chapters

The seven following chapters explore the space of autographic design through many examples. The chapters are loosely organized around design interventions of all kinds: found traces to the design of instruments, self-registration, and analog simulations. My focus also includes designed materials and materiality in digital systems. Each chapter provides historical and theoretical context and explores the linkages between autographic design and data visualization practices. While the subject raises many conceptual questions, I intend this to be a practical book that encourages experimentation.

Chapter 1 discusses the theoretical foundations of autographic design. The trace is a notoriously polyvalent concept that is hard to pin down—a semantic field with sometimes conflicting implications. Traces exist in an epistemic state of in-between: they are features of the world that at the same time emerge in the eye of the beholder.[68] The chapter briefly examines how different schools of thought, from semiotic to presentational and relational theories, address these tensions.

Chapter 2 then summarizes the operations of autographic design, which form the building blocks for autographic systems. Many fields of practice have developed methods for revealing traces. Their methods are usually considered in isolation and are rarely discussed outside their discipline since the steps involved in data production are usually of secondary interest outside their domain. This chapter takes the broad spectrum of these trace-making practices and seeks commonalities and general principles.

Chapter 3 is dedicated to environmental traces. These are mostly unintended by-products of various processes. Their discovery requires cultivating skills of observation. These traces are used to reconstruct past events or navigate an environment. Architects and planners use such traces as unobtrusive measures of human behavior in space: what people do, where they gather, and how they modify their surroundings. While patterns of wear and accretion are ubiquitous in the lived environment, they are also popular metaphors in digital interfaces. The chapter makes the seemingly paradoxical claim that trace-reading is a form of design intervention—not just because there are practices in which, for instance, pictures of vehicular traces are annotated and circulated by bicycle activists, or where skid marks are marked by accident reconstructionists, but also because observation involves the active probing and tuning into a phenomenon, process, or event.

Chapter 4 focuses on instrumentation, using the atmosphere as a case study. Weather is a sensory phenomenon that can be felt with the whole body but is nevertheless difficult to assess without the use of instruments. The chapter investigates autographic approaches to measuring atmospheric phenomena, ranging from sensory judgment to complex instruments. While instrumentation shifts the understanding of climate from a sensory to a statistical phenomenon, bioindicators, sentinel species, and other proxy data sources can make environmental data experiential. The chapter examines the story of how environmental activists and grassroots scientists have mobilized various kinds of sensory experiences, from the human to the nonhuman, to contest the picture offered by official data records.

Chapter 5 focuses on methods of self-inscription to create and preserve traces, as well as how trace-phenomena are operationalized to construct evidence. While trace-making practices in the sciences are concerned with generalization, forensic materiality focuses on individualization.[69] Starting with methods of self-registration and nonocular photography, the chapter examines the rhetorics of the trace and how it is employed by citizen scientists and amateur forensic experts to support truth claims. Persuasive techniques in data visualization have an equivalent in autographic design: both involve framing, but while the former represses the materiality of its sources as a matter of method, the latter involves making overt the contextualization of traces.

Chapter 6 is dedicated to analog visualization systems that I describe as “autographic environments.” They are composed through various design operations and create an environment isolated from external influences, in which a phenomenon can unfold. A familiar example of an autographic environment is the wind tunnel, an isolated space that involves various means of tracing such as smoke, tufts, and clay. Autographic environments are analog computers that make otherwise singular events repeatable. They provide a platform for performing the same visualization tasks using different inputs, and for observing the results. Since autographic environments are analog computers that can solve a specific problem, it is no coincidence that the mathematician John von Neumann worked on wind tunnels before he outlined the basis for the modern computer. Autographic environments are often hybrid environments that include digital and analog components that are coupled in various ways.

Chapter 7 focuses on autographic data and the conditions under which digital data can be examined as a set of physical traces. Deepfakes and generative adversarial networks (GANs) have given rise to new levels of disinformation. Focusing on the concept of the signal, this chapter examines the interface between autographic practices and the current digital modes of visualization and analysis. It examines the various ways in which underlying material infrastructures manifest themselves in digital data as unintended and even unwanted traces, as well as how different practices use these "material residues" in data as an entry point for critical analysis.

The conclusion of this book takes a brief look at the relevance of autographic design in emerging areas such as metamaterials and synthetic biology, before concluding with a set of recommendations on how autographic perspective can be applied to working with data.

While a significant part of STS literature has examined practices of trace-making and visualization in the context of scientific experimentalism, this book highlights the arts and do-it-yourself culture, including self-reflective artists, community scientists, and amateur forensic experts. There are several reasons for putting these groups in the spotlight. First, these practices demonstrate the close epistemic and cultural links between science and the public sphere. Academic disciplines outside the realm of so-called pure science, such as epidemiology and environmental justice, depend on the involvement of the public in many ways, including the articulation of problems and the implementation of research design. Second, I hope to show that public science in its many forms offers hints about a new modality of research. Rather than relying on an analytic mode that involves isolating and dissecting a phenomenon into discrete components, amateur practices consider the phenomenon as a whole, with all its social and material implications. Their approach to research is relational—not just in the sense of material but also human relationships. The holistic dimension of the network is lost when we study it solely by separating it into its components. We discover this dimension by characterizing its behavior as a whole. By probing, intervening, and collectively reflecting on the effects of various kinds of trace- and object-making, these do-it-yourself practices constitute a research modality that harks back to the empiricist tradition of the public experiment.