A PLACE FOR PLANT DATA

ACC_NUM, HABIT, HABIT_FULL, NAME_NUM, NAME, ABBREV_NAME, COMMON_NAME_PRIMARY, GENUS, FAMILY, FAMILY_COMMON_NAME_PRIMARY, APG_ORDER, LIN_NUM, ACC_DT, ACC_YR, RECD_HOW, RECD_NOTES,

PROV_TYPE, PROV_TYPE_FULL, PSOURCE_LABEL_ONE_LINE, COLLECTOR, COLL_ID, COLLECTED_WITH, COUNTRY_FULL, SUB_CNT1, SUB_CNT2, SUB_CNT3, LOCALITY, LAT_DEGREE, LAT_MINUTE, LAT_SECOND, LAT_DIR, LONG_DEGREE, LONG_MINUTE, LONG_SECOND, LONG_DIR, ALTITUDE, ALTITUDE_UNIT, DESCRIPTION, COLLECTION_MISC

75–67, T, TREE, 617, AILANTHUS GIRALDII, AIL. GIRALDII, GIRALD, AILANTHUS, AILANTHUS SIMAROUBACEAE, QUASSIA FAMILY, SAPINDALES, 75–67, 25 JAN 1967, 1967, SG, U, UNCERTAIN, MR. THOMAS ELIAS, U. S. NATIONAL ARBORETUM, 3501 N. YORK AVENUE, N.E., WASHINGTON, DC 20002 *VIA*, BOTANICAL GARDEN OF THE UZBEC ACADEMY OF SCIENCE, DSHACHAN ABIDOVI, 232, TASCHKENT, UZBEC SSR, COMMONWEALTH OF INDEPENDENT STATES

***Source*: Arnold Arboretum**[1]

ASSEMBLING A VIEW FROM DATA

On a clear day in 2014, a small group of colleagues and I set out to photograph Bussey Brook Meadow from above, using a digital camera tethered to a weather balloon.[2] The meadow, a small stretch of land located in the Boston neighborhood of Jamaica Plain, is an experiment in urban ecology. It is also part of the Arnold Arboretum: a research collection of plants, vines, and shrubs managed by Harvard University. Bussey Brook was established more than twenty years ago to learn what wild varieties might flourish when left unattended in the contemporary environmental and social conditions of the city.[3] The answer, in part, is in our photographs: a lush composite of native and nonnative species, with the sky-seeking "tree of heaven" (*Ailanthus altissima*) dominating the tree canopy (figure 2.1).[4]

Using low-tech instructions for aerial photography developed by the Public Laboratory for Open Science and Technology (Public Lab for short), we made a harness for a point-and-shoot camera out of an empty soda bottle and secured it to an inexpensive but rugged helium-filled balloon.[5] The camera was set to continuous capture. As our rig ascended, it took more than a thousand photographs, each successive shot framing an expanded view of the meadow below.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

2.1

Selected photos from more than a thousand aerial images of the Arnold Arboretum's Bussey Brook Meadow created with a balloon camera. Image by the author in collaboration with metaLAB.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

The first few photographs captured only our small group, huddled near the end of a nylon rope trailing behind the balloon. But in the subsequent images, these tight boundaries expanded, encompassing larger and larger swaths of green at each fraction of a second, until the balloon was at its final height of about five hundred feet. At this height, the frame encompassed many things: the taut string tethering the balloon, our team dotting the ground below, a tangle of plants unidentifiable to all but the most expert observer, and a path weaving between banks of foliage. As the camera pivoted with the wind, its frame shifted again to include the extents of the meadow, adjacent arboretum grounds, and Oak Grove train station, which marked our place at the western edge of Boston.

These images aided arboretum staff in understanding and evaluating their urban ecology experiment. The documentation was necessary because Bussey Brook Meadow is the only part of the arboretum that is not carefully curated and cataloged. As such, our photographs helped to shape the overarching project. But more than a collection of plants to be identified discretely, our images revealed the meadow as a place: our presence, the path, and the proximity to the T stop (Bostonian for "train station") give this feral stand of trees an identity that transcends the sum of its parts.[6]

My interest is not in aerial photography per se but instead in understanding how varied forms of data—of which photographs can be an example—shape places, and reciprocally, how those places shape their data. Beginning with photographs of the arboretum helps us think about what's missing from our understanding of another kind of data: the accessions records of the arboretum's curated collection. We are used to identifying places through images, such as those taken over the meadow. But we cannot say the same for other forms of data. What does it mean to see a place through collections data?

ENCOUNTERS WITH DATA: ABOUT, IN, AND FROM A PLACE

Data have complex attachments to place, which invisibly structure their form and interpretation. This is the second of six principles that frame the book. Place is routinely overlooked as a dimension of situatedness in social studies of data. As I remarked in chapter 1, often but not always, *situated* refers to embodiment or social context. In this chapter, I make use of the arboretum's accessions records to illustrate the manifold relationships between data and place.

Established in 1872 and located on 281 acres, the Arnold Arboretum is equal parts urban laboratory and "zoo for plants" (figure 2.2). It is one of the most comprehensive, well-documented collections of its kind in the world.[7] Its accessions records are one of many genres of data in use at the institution. I have chosen to work with these records rather than other forms of data collected at the arboretum, introduced later in this chapter, because of the evocative ways in which accessions details highlight important place attachments that data can hold.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

2.2

Map of the Arnold Arboretum. Bussey Brook Meadow is on the right. Courtesy of the Arnold Arboretum.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

Hosting around fifteen thousand living plants today, and about seventy thousand over the course of its history, the arboretum is an apt site for investigating data's attachments to place for four distinct reasons. First, place is a key piece of data for the arboretum. Its collections are assembled from sites of scientific and cultural significance around the world. Second, the arboretum has long sought to be a place in which scientists and citizens alike can encounter large collections of data firsthand simply by walking the landscape, and discovering a variety of carefully tagged trees, vines, and shrubs. Third, when understood as a set of conditions for production, place has shaped fluctuations in botanical data over the course of the arboretum's long history. Fourth, when looked at graphically, the arboretum's data can help us see place in new ways, which aren't limited to aspects of geolocation.

In summary, data can be about place, in place, from place, and even generative of place. Learning about these long-standing forms of place attachment can prompt us to challenge settled conceptions about the relationship between data and place in contemporary life. Each of these attachments to place is a different way in which data are subject to local examination. The dimensions of place attachment identified in this chapter along with the means of identifying them suggest a place-based approach to understanding initially unfamiliar data sets in terms of their settings. Let us start by taking a look at how local readings of data related to individual specimens might reveal diverse place attachments.

Place as Data: The Case of *Prunus Sargenti*

> PROV_TYPE, PROV_TYPE_FULL, PSOURCE_LABEL_ONE_LINE, COUNTRY_FULL, SUB_CNT1, SUB_CNT2, SUB_CNT3, LOCALITY, LAT_DEGREE, LAT_MINUTE, LAT_SECOND, LAT_DIR, LONG_ DEGREE, LONG_MINUTE, LONG_SECOND, LONG_DIR, ALTITUDE, ALTITUDE_UNIT

The fields listed above (beginning with PROV_TYPE and ending with ALTITUDE_UNIT) all contribute to the characterization of place in arboretum accessions data. In order to understand the origin of a single specimen using these data, it is necessary to take account of multiple fields and how they might be interrelated. Already mentioned in the introduction, a special cherry tree (*Prunus sargentii*) accessioned to the arboretum on a leap day in 1940 provides an example of this process. The history of the tree is cataloged in a custom digital record system called BG-Base, under the specimen number 130–40. The provenance of the plant (PSOURCE_LABEL_ONE_LINE) is attributed to the institution's founding director, Charles Sprague Sargent, at the address of the arboretum itself: "125 The Arborway, Jamaica Plain, MA." Part of the tree's Latin name, *sargentii*, honors this parentage.[8] Meanwhile, its country of origin (COUNTRY_FULL) is listed as "Japan." Sargent might have acquired the plant during an expedition to Asia. But this would seem to be in conflict with other known conditions—first and foremost,

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

that Sargent died in 1927, thirteen years before the listed accession date. Moreover, for reasons that will be explained later, wild plants from abroad had not been taken in at the arboretum since the mid-1920s.

Sargent could not have transported the plant from Japan to the arboretum on the date of accession. This apparent inconsistency is an artifact of the way that origins are documented at the long-lived institution. Staff at the arboretum know that a few select fields—date, place of origin, and provenance—don't tell the whole story. One has to look to another field, the provenance type (PROV_TYPE) of the cherry tree, to learn that it is a "cultivated plant of known (indirect) wild origin" or "Z" for short. In other words, specimen number 130-40 grew from a cutting taken off another plant collected much earlier. Provenance type is a classification of disputed value, for it is a social distinction versus a biological one. Wild plants and their cuttings are genetically identical. In this case, the "Z" helps to clarify that the cherry tree in question was grown from a cutting of one of Sargent's original specimens—probably number 16760, unearthed from its native Japanese soil in 1892. The apparent difference in the formats of these two specimen numbers is just another indication of how data creation practices can change over the life of an institution. The former (130-40) has a two-digit year attached (the part after the dash) to indicate that the specimen was accessioned in 1940. The latter (16760) was created before that practice was adopted.

This example illustrates some of the complexities of place as presented within the arboretum's data. Data *about* place are not simply contained in a field. This form of place attachment must be understood through a matrix of values, coordinated through local knowledge about the history of data collection practices and how they encode place as a subject.

Place of Data: The Case of *Torreya Grandis*

As the previous instance shows, the arboretum is an aggregated landscape stitched together from plants once residing in other places. Most plants hail from ecological zones similar to that of Boston's, stretching across England, Greece, South Korea, China, and Japan. When encountered at the arboretum, each of these plants stands with its data. A thin plastic card embossed with a subset of accession details hangs from its trunk or branches (figure 2.3). The cards contain fields that are relevant for arboretum staff, researchers, and visitors: scientific name, accession number, plant family, accession date, propagation material (e.g., seed "SD" or scion "SC"), location, common name, and source/collection data. Together the plants and their tags transform the arboretum into a full-scale scientific map organized using the Bentham and Hooker taxonomy—a system that dates to the late nineteenth century. The arboretum landscape is itself a place for encounters with data.

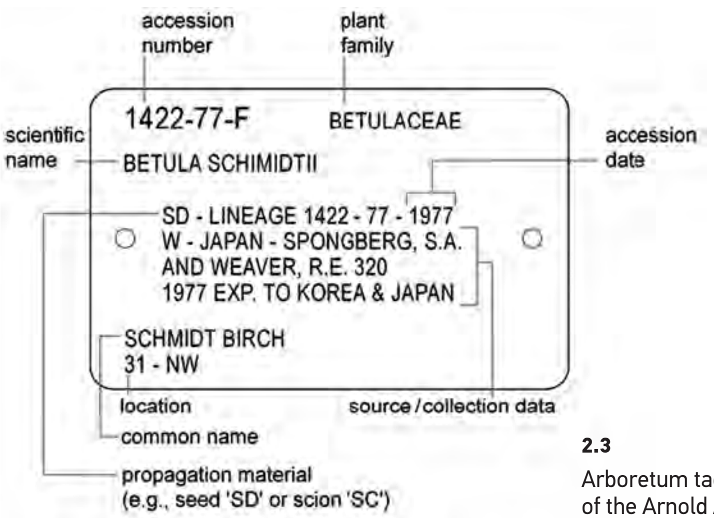
In order to understand this second form of place attachment, let us revisit a tour of the grounds that occurred in late June 2013. During a workshop that I coorganized, a group of visitors were guided by arboretum senior researcher Peter Del Tredici

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

through the Explorer's Garden, an area nestled in a microclimate beneath the summit of Bussey Hill. Del Tredici stopped to comment on his relationship to the living collections. "I've got a lot of direct connection to a lot of these plants. That little plant, *Torreya grandis*, I collected in China in 1989. So a lot of these are like my offspring."[9] Del Tredici explains that he found the seeds of the *Torreya grandis* at a market in China. Fleshy and green, they struck him as unusual examples of edible seeds produced by a conifer. But beyond what is interesting about the plant itself, this quote provides a compelling starting point for understanding what is and is not included in the information landscape of the arboretum.

The acquisition date of the *Torreya grandis* and Del Tredici's association with it are duly noted on the plant's tag. Also pressed into the tag's smooth surface, the term *pinales* registers the plant's status as a conifer. There is no hint of what Del Tredici has referred to as the "oddness" of this ordering.[10] Indeed, several features of the plant's local significance are not included on the tag, which serves mainly to position the *Torreya grandis* within a scientific landscape. Tags do not explain how plants like this one are literally and figuratively torn up by the roots, and then relocated to a new ecological and cultural context. Let's explore a few of the conditions that don't count as data in this context.

Del Tredici is identified on the tag as a "collector," not as "progenitor" or "breeder," as his statement would suggest—this, despite the fact that he is responsible for the reproduction of the plant in the Boston region. The term *collector* speaks of the scientist-and-specimen relationship between Del Tredici and the plant, rather than the more nurturing association between Del Tredici the horticulturalist and the organism he has cultivated. The latter is more in line with his own intimate way of identifying the *Torreya grandis* as his "offspring."



2.3

Arboretum tag diagram. Courtesy of the Arnold Arboretum.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

More generally, there are few traces of the fruitful intersections between the living collections and local communities in Boston that surround the arboretum. Don't look to data for connections between "dandelions" (*Taraxacum officinale*) and the elderly Greek women who collect them from the arboretum grounds in the early summer to make *horta vrasta* (boiled greens), or associations between the "tree of heaven" (*Ailanthus altissima*) and the devout Dominicans who discover starlit sites for their Santeria rituals in the groves of Bussey Brook Meadow. Such details, though important to the local meaning of the arboretum's plants, are not part of the way that data on tags interact with the place.

I introduce the example of the *Torreya grandis* to call attention to the placement of data, but also their limits as tools for understanding the places in which they reside. While useful as a means of establishing shared references among the arboretum's staff and its visitors, data do not capture the full lives of arboretum plants.[11] Data categories sit beside yet do not account for all the varied place-based meanings that the plants embody.

Data of Place: The Case of *Tsuga Caroliniana* and *Tsuga Canadensis*

So far I have revealed how place appears in data, and how data appear in place. There is also the important matter of how a place affects data's production. For this last point, let us consider the Carolina and eastern (or Canadian) hemlocks (*Tsuga caroliniana* and *Tsuga canadensis*), which are trees local to the East Coast of the United States. Both have been in rapid decline due to a nonnative insect, the hemlock woolly adelgid (*Adelges tsugae*). In April 1997, an unaccessioned stand of almost two thousand hemlocks on Hemlock Hill—originally brought to the arboretum many years earlier, not as scientific specimens, but as filler plants meant to occupy a bald spot on the landscape created by a destructive hurricane—fell victim to the pest. A note in the accession record for one Carolina hemlock reads, "Plants producing very heavy seed crop, heavily infested with woolly adelgid."[12] Over the winter of 1997–1998, the trees were "labeled, mapped, and qualitatively assessed" to monitor damage caused by the infestation.[13] Although these trees had been residents on the institution's grounds for decades, they were only accessioned into the collection in order for the infestation to be tracked and treated with imidacloprid, a powerful insecticide. The hemlocks were never meant to be an official part of the collection. Regardless, the accession of the blighted hemlocks made 1997–1998 a peak moment of expansion for the arboretum, but only from the perspective of data.

This example demonstrates that even seemingly straightforward fields like "date" can have a complex relationship to place. For each entry in BG-Base, what the accession date means is dependent on local conditions. It might mean when a seed was planted, when a seedling arrived on site, or simply—as in the case of these hemlocks—when an existing plant was annexed to the collection. But beyond the curious and local significance of their accession dates, the hemlocks are interesting because they raise deeper issues about the role that data perform at the arboretum.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

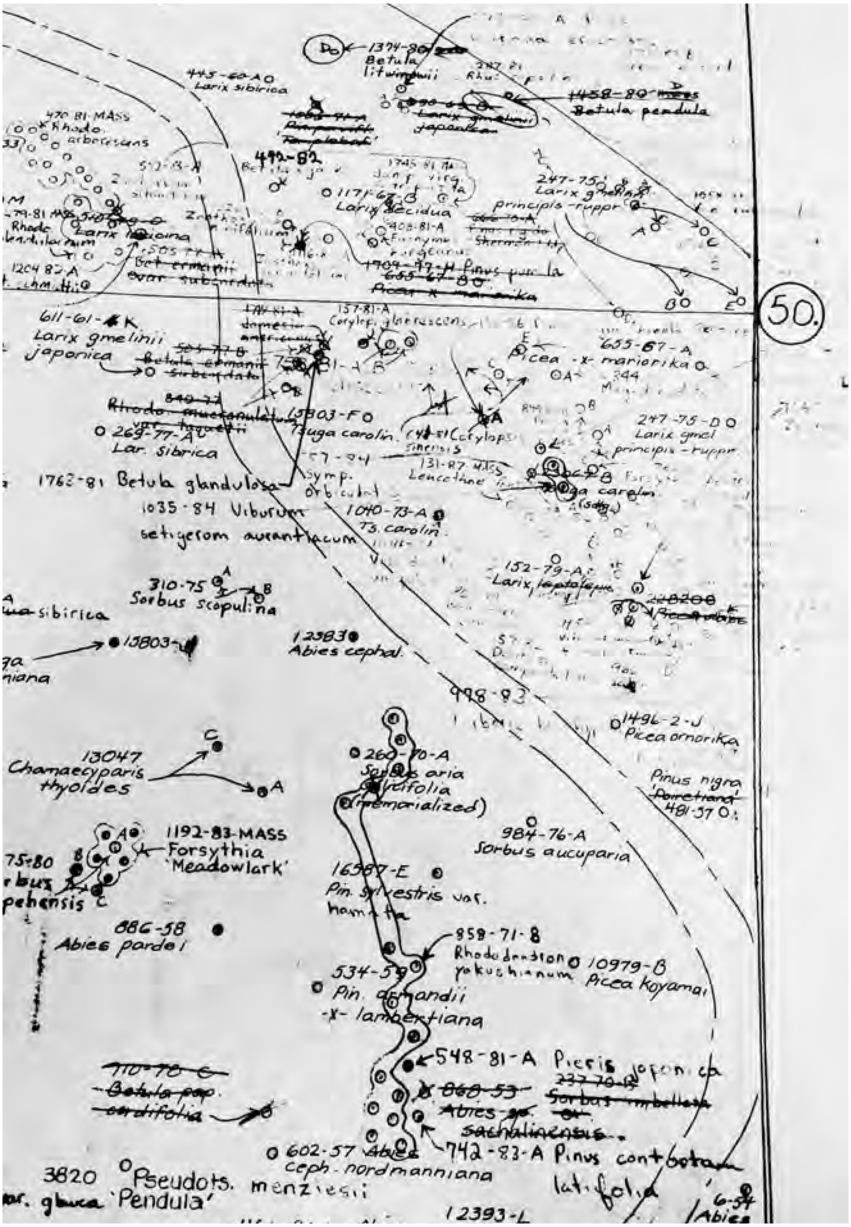
Controversy still surrounds the decision to make the stand of hemlocks part of the collection. Del Tredici, who originally argued for their accession, continues to see the trees as invaluable for the study of the infestation process. "It was only by accessioning the plants that we could track their decline over time or the insecticidal treatment."[14] Meanwhile, current director William Friedman, who arrived years after the hemlocks were accessioned, looks on these trees of questionable provenance as inherently undesirable, for they lack essential data about their origins that would make them reliable subjects of scientific study. Why not replace them with trees of actual research significance?

Such disagreements highlight the tensions between competing realities at the arboretum: it is a living place, but also a repository for data. Hence data may be looked on as "just good enough" to support the care of the collection: organizing plants, notes, and relationships among them in a convenient manner.[15] But without reliable data, the emergent form of the collection can disappear altogether, with its contents scattered in an ontological wild.

Coexisting concerns about the necessity of data and their inherent instability over time reinforce a lesson from STS that holds across shifts in technology: data must be part of a knowledge system, or what Paul Edwards calls a "knowledge ecology."[16] The connection to environmental processes is apt. Arboretum scientists, specimens, and information infrastructures are all necessary to generate, verify, and sustain what the place knows. It is the encompassing place—of which data are only a part, along with the people and plants—that holds knowledge about the arboretum hemlocks, their deadly infestation, and its implications for similar trees across the Northeast. At the arboretum,

2.4
A hemlock tree at the arboretum.
Image by the author.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023



2.5

Early map of the arboretum.
Image by the author.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

the knowledge ecology is more than a metaphor. Data are necessary components of the functioning ecology created and maintained there. The story of the hemlock trees (figure 2.4) illustrates how becoming data can be a prerequisite for receiving sustaining care. Thus data can transcend their roles as representations by directly supporting the places that they describe.

READING DATA IN PLACE

In order to create these local readings of accessions data, I have relied on a prolonged engagement with the Arnold Arboretum. During the period from 2012 to 2014, I lived and worked in close proximity to the institution in Jamaica Plain. I conducted interviews with researchers, administrators, and technologists, and searched through archival sources at their library. More important, though, I was a participant observer in both formal and informal engagements, including a course on landscape architecture taught by Del Tredici, the series of outings to photograph Bussey Brook Meadow, and the multi-day workshop that I coorganized to bring together arboretum staff with scholars of science and technology. Over the course of the final year of this engagement, I worked with the staff to develop reading techniques appropriate for looking at their data.

Seeing data as texts accessible to traditions of hermeneutic inquiry means reading them within an interpretative context. It would be difficult to understand these records without considering their historical attachments to the arboretum as a place. Indeed, its accessions records have a long history of development and use. For one thing, they were not always recognizable as data. The arboretum has weathered many successive regimes of documentation (figures 2.5–2.7).

Each organism at the institution has germinated within a social and technological setting, its care and curation managed through the instruments and information structures deployed during its lifetime. These place-based practices along with the documents that they produce register what is valued about individual organisms and, in turn, how those values change over time.

Today, plants collected from around the world and across time are held together by BG-Base. Each entry in the arboretum's data set includes an accession number, an extensive list of scientific, common, and abbreviated names, redundant ways of identifying the time of accession, the form and mechanism of reception, individuals associated with the plant, various descriptions of the place that the accession hails from, its condition in the wild, and an additional catchall category. A list of fields used by the arboretum includes:

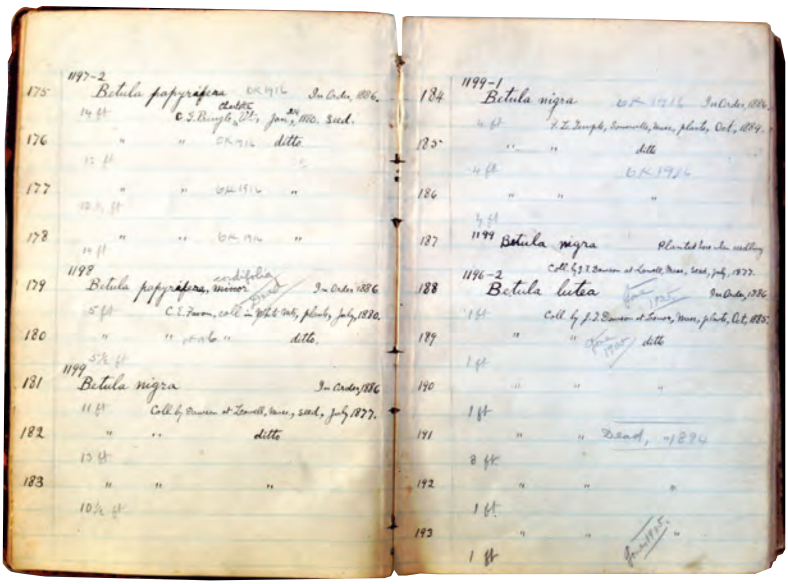
> ACC_NUM, HABIT, HABIT_FULL, NAME_NUM, NAME, ABBREV_NAME, COMMON_NAME_PRIMARY, GENUS, FAMILY, FAMILY_COMMON_NAME_PRIMARY, APG_ORDER, LIN_NUM, ACC_DT, ACC_YR, RECD_HOW, RECD_NOTES, PROV_TYPE, PROV_TYPE_FULL, PSOURCE_LABEL_ONE_LINE, COLLECTOR, COLL_ID, COLLECTED_WITH, COUNTRY_FULL, SUB_CNT1,

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

SUB_CNT2, SUB_CNT3, LOCALITY, LAT_DEGREE, LAT_MINUTE, LAT_SECOND, LAT_DIR, LONG_DEGREE, LONG_MINUTE, LONG_SECOND, LONG_DIR, ALTITUDE, ALTITUDE_UNIT, DESCRIPTION, COLLECTION_MISC

If encountered within a library, museum, or archive, many of these fields would be considered metadata: the information necessary to catalog a book or other object, such as details of their contents, context, quality, structure, and accessibility. At the arboretum, this locally defined selection of fields is known simply as "accessions data." But they are shaped by many of the same forces that affect metadata.[17] For example, each accession record exists as part of a local constellation of information, including the details of the associated plant's phenology, genetic characteristics, transpiration rate, and growth habit. Even the specimen itself is a kind of data.[18] This entire "data assemblage" is necessary to make plants real as well as present in the contemporary ecological, scientific, and public life of the arboretum.[19]

As mentioned above, documentation practices at the arboretum long predate contemporary notions of data. Today, records are available in multiple formats simultaneously: on maps (figure 2.5), in ledgers (figure 2.6), on index cards (figure 2.7), and only recently, in digital form. It wasn't until summer 1985 that the arboretum started converting its accessions data from index cards crowded in a vertical file to digital data stored in BG-Base. These digitized data afford new opportunities for access and



2.6

Early ledger containing accessions information. Image by the author.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

analysis. Even so, some staff members continue to use older formats exclusively for they do not yet trust the process of digitization or interpretations of outsiders with new-found access to their data.

Regardless of the format, what counts as data at the arboretum is a matter of context. As Del Tredici explains, "The data, in and of itself, is only valuable [for] somebody who understands its significance." To further his point—one that I have tried to echo throughout the book—Del Tredici likens the "raw data" to seeds. When a seed won't germinate, there are innumerable possible reasons. "Unless you know how to interpret the behavior of the seed, it is just nondata."[20]

VISUALIZING PLACE

Through my local readings of the Arnold Arboretum's accessions records, I have sought to reveal numerous ways in which data can be entangled with place: when place is a kind of data, when place is the site of encounters with data, and when place is the site of data's production. Each of these place attachments can be exposed through local readings of accessions data for particular plants: *Prunus sargentii*, *Torreya grandis*, and *Tsuga caroliniana*. These case studies were triggered by discrete technical problems or controversies that I happened upon. As such, they provide a kind of event-based reading of data. Yet looking at the accessions of the arboretum altogether through visualization techniques can reveal alternative conceptions of place.[21]

2.7

Card catalog containing accessions data before it was digitized. Image by the author.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

I use the term *visualization* to describe the experience of looking at the whole arboretum through the data of its parts. Rather than being a god's-eye view, characterized by theorist of science and technology Donna Haraway as one that seems to come "from nowhere, from simplicity," these visualizations offer situated but wide-ranging perspectives: views of data as opposed to views through data.[22] Creating this kind of visualization requires a critical sensibility toward data, including attention to what might be occluded and what other vantage points are possible. This approach complements prior work in geography on the critical studies of landscape representation as well as the development of critical practices in mapping.[23]

My visualizations are meant to function more like panoramas than maps. Although it is not uncommon to hear the term *panorama* used today to describe graphical displays of data, few acknowledge that unlike maps, panoramas are situated ways of seeing places. As far back as the eighteenth century, the term was used to describe pictorial representations of landscapes as seen by an observer positioned at a single strategic point. Moreover, panoramas have long been understood as mediated. Like visualizations, they are enacted through technological means. For instance, historian Wolfgang Schivelbusch uses the term *panoramic* to evoke the once-unfamiliar view across an expansive landscape afforded by the speed of the passenger train.[24] Just as the rapid pace of the passenger locomotive offered new vistas across broad stretches of space, the visualizations included here reveal perspectives at previously incomprehensible scales. Visualizations aren't narrowly defined technical tools; they generate alternative experiences of data and the places that they depict.[25]

In the visualization presented below (figure 2.8), the arboretum is portrayed as an aggregate, pattern, and system in flux. Here, data are enlisted to construct a new sense of place. Because of their scale and heterogeneity, large digitized data sets offer opportunities for experiences of place that—like Schivelbusch's locomotive panorama—are different from anything seen before.

In the next section, I introduce a series of experimental visualizations. None of these are neutral or inevitable. Rather, they help us reimagine the arboretum as a place with origins, structures, and dynamics that are not merely geographic.

Place as History

Figure 2.8 portrays the arboretum as an aggregate place developed over time. Beginning in 1872 and ending in 2012 (when this set of records was made available for use), the visualization portrays a temporal graph of plant specimens. The image is a kind of timeline, structured by yearly accessions, much like trees record environmental changes in their annual growth rings. Months and days index accumulated plants, each denoted by a dot. This two-dimensional view can be enhanced by a series of section cuts through daily accessions (figure 2.9). In the original interactive version of this visualization, the section cut, which portrays the number of accessions on each day of the selected year, can be produced for any year along the timeline.[26]

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

Such visualizations can be used to call attention to variations in the data by linking them to color, size, and other visual cues. For instance, figure 2.8 displays changes in provenance type (a category mentioned earlier) across the history of the collection. Here a green dot represents a plant collected in the wild, a yellow dot signifies a cutting from a wild plant, a black dot indicates a cultivated plant, and a gray dot stands in for a plant from an unknown origin (far more common than one might expect). Fluctuations across these provenance-related colors illustrate shifts in the makeup of the arboretum from collections of scientific importance (mostly collected from the wild) to selections in the service of horticulture (mostly from other cultivated collections).

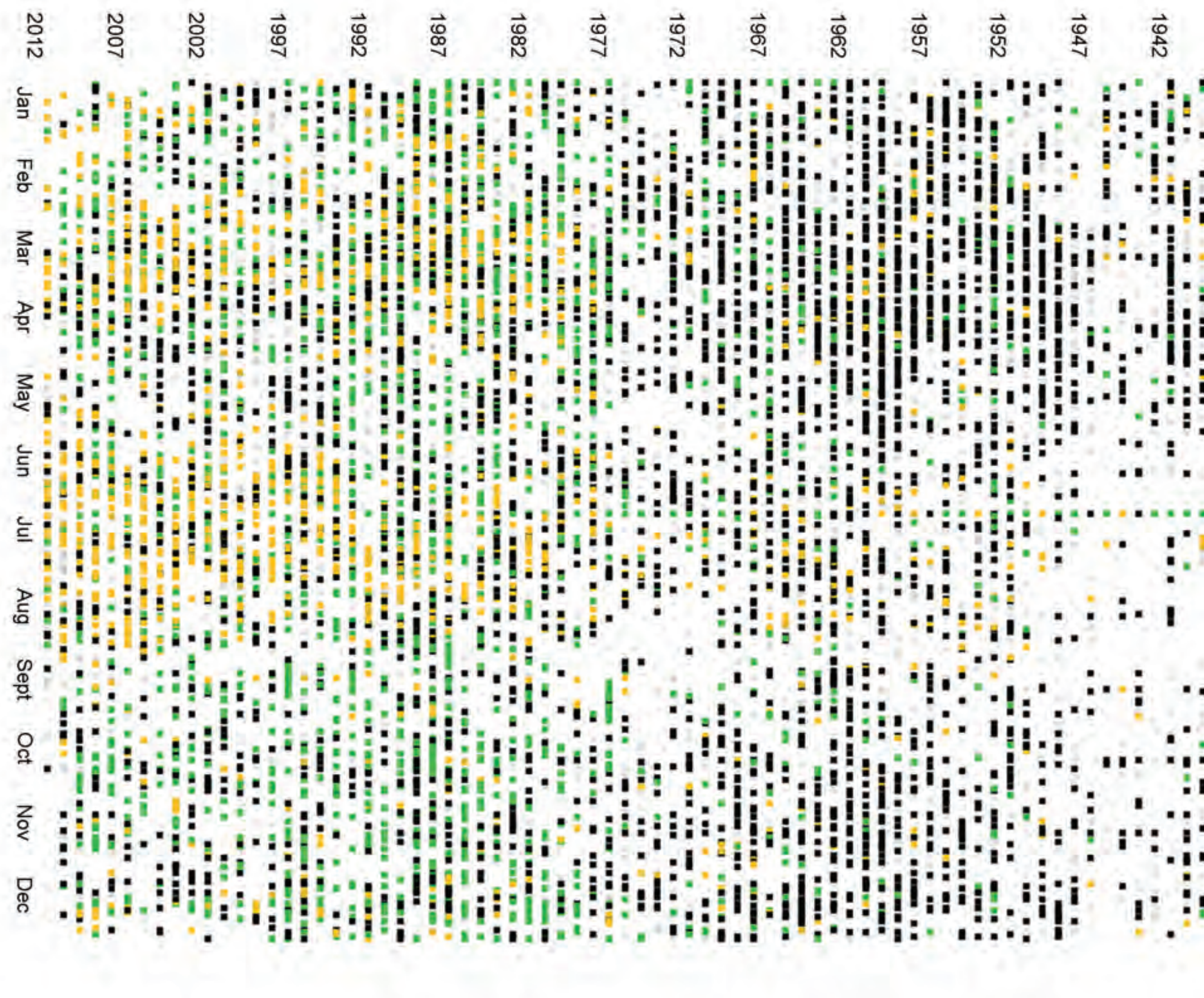
The distribution of green, yellow, black, and gray dots faintly demarcates three eras of collecting identified by curator of living collections Michael Dosmann.[27] In the late nineteenth and early twentieth centuries, Sargent engaged in a global project of scientific fieldwork to collect distantly related species from around the world as evidence to support Charles Darwin's theory of evolution. In the 1920s, however, the US Department of Agriculture discovered that the arboretum was inadvertently collecting invasive bugs along with its imported plants and took action to stop it. Sargent lost the ensuing legal battle, and wild collecting decreased substantially thereafter.

The middle years of the twentieth century are sometimes referenced by staff as the Wyman era, after a prominent horticulturist. During this time, the arboretum halted its foreign expeditions and relocated its scientific research to Harvard's Cambridge campus. The next phase of research centered on the herbarium, a much larger collection made up entirely of dried plants (figure 2.10). Dosmann explains that the expansive grounds in Jamaica Plain became a "showcase garden," a place to display the horticultural trends of the day. In this period, says Dosmann, "if you did want to go and collect anything, you went to a nursery."[28]

It wasn't until the early 1970s, during a reevaluation of the mission of the collection, associated with its centennial, that the arboretum reinitiated its expedition work abroad. The renewal of overseas fieldwork expanded relationships with institutions in Asia, and later, studies on emergent and imperative questions around global climate change. My provenance visualization registers some aspects of these long-term temporal shifts, highlighting in particular the relationship between the two defining arms of the arboretum, scholarship and horticulture, and the ways in which their relationship changed over time.

This visually oriented reading of the arboretum as data is unlike a photograph or geographic map of the place. It highlights a landscape shaped over time by otherwise-invisible ecological, organizational, and even political forces. This particular use of the method is but one way of reading. The data support many alternative portrayals of place.

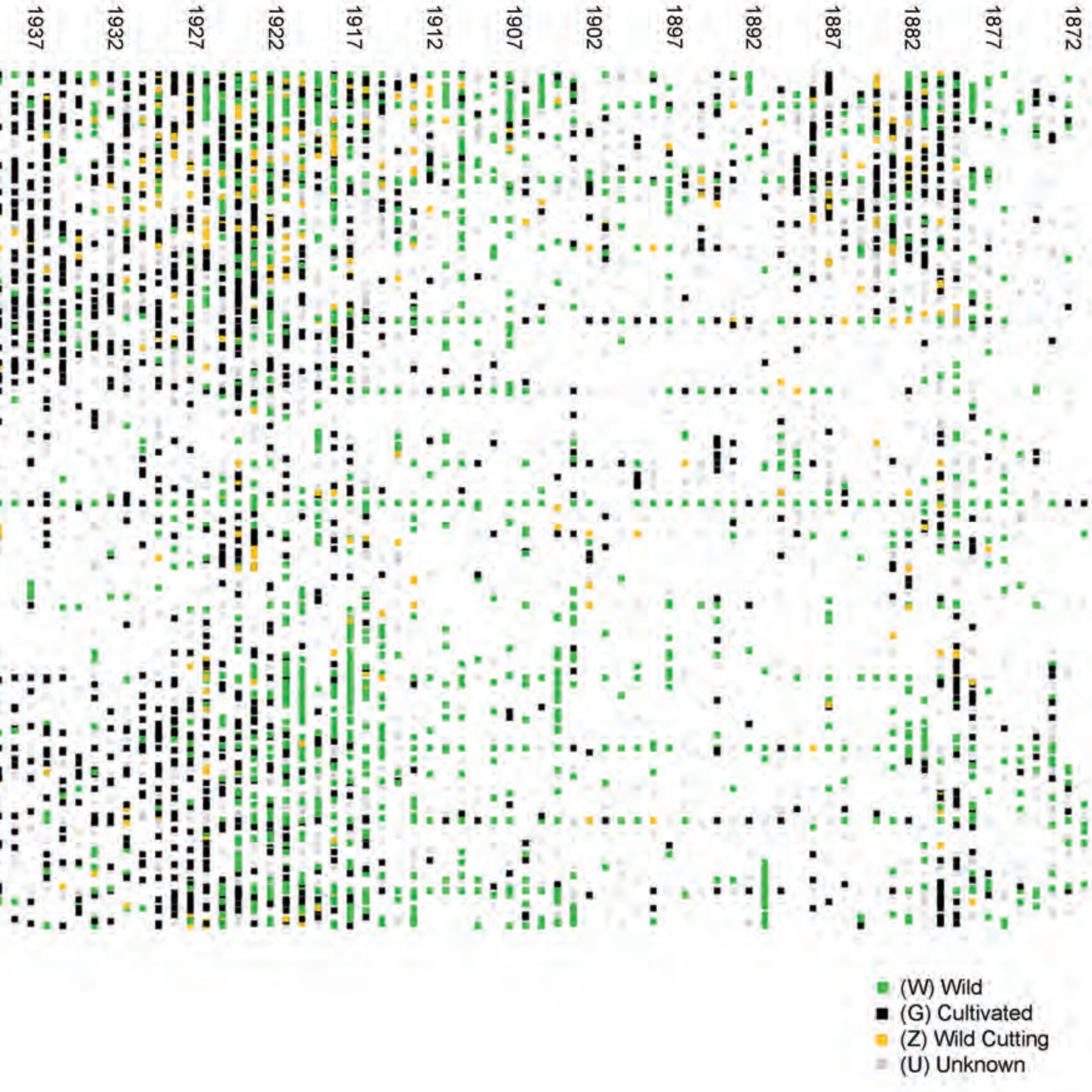
Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023



2.8

Linear timeline of arboretum accessions. Plants without accession dates, of which there are 1,190, are not included here. Image by the author and Krystelle Denis.

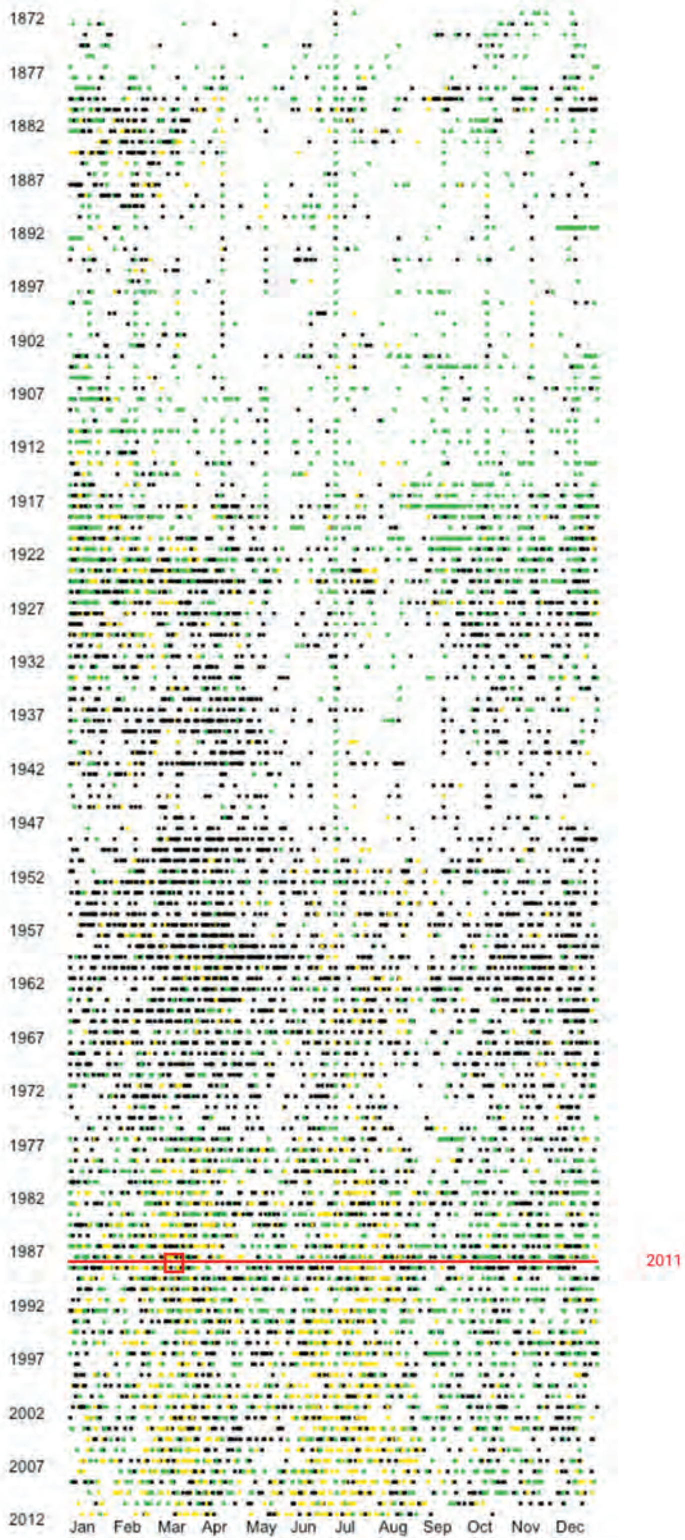
Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023



2.9 (following pages)

Section cut through linear timeline of arboretum accessions. Image by the author and Krystelle Denis.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023



Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

8 ACCESSIONS ON 8-MAR-11

44-2011: MAGNOLIA X SOULANGEANA, SHRUB/TREE, MAGNOLIA FAMILY, U. S. NATIONAL ARBORETUM, 3501 NEW YORK AVENUE, WASHINGTON, DC 20002, COLLECTED BY: UNKNOWN

45-2011: SYRINGA PEKINENSIS 'MORTON', CULTIVAR OF LILAC, TREE, OLIVE FAMILY, KNIGHT HOLLOW NURSERY, INC., 7911 FORSYTHIA COURT, MIDDLETON, WI 53562, COLLECTED BY: UNKNOWN

46-2011: CERCIS CANADENSIS "SPORT", BEAN FAMILY, ARNOLD ARBORETUM, 125 THE ARBORWAY, JAMAICA PLAIN, MA 02130, COLLECTED BY: UNKNOWN

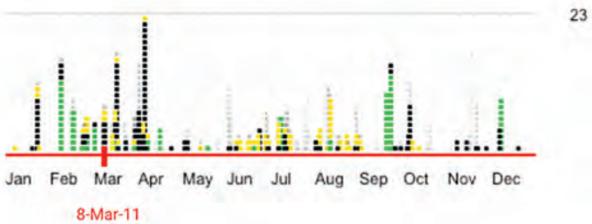
47-2011: CERCIS CANADENSIS, EASTERN REDBUD, TREE, BEAN FAMILY, MR. MARK KRAUTMANN, HERITAGE SEEDLINGS NURSERY, 4194 71ST AVENUE SE, SALEM, OR 97301-9242, COLLECTED BY: UNKNOWN

48-2011: MALUS FUSCA, OREGON CRABAPPLE, SHRUB/TREE, ROSE FAMILY, UNIVERSITY OF BRITISH COLUMBIA BOTANICAL GARDEN, 6501 NORTHWEST MARINE DRIVE, VANCOUVER, BRITISH COLUMBIA V6T 1W5, CANADA, COLLECTED BY: MACPHAIL, J.

49-2011: MALUS SP., SHRUB/TREE, ROSE FAMILY, HORTICULTURE DEPT., MICHIGAN STATE UNIVERSITY, EAST LANSING, MI, COLLECTED BY: BARON, M.

77-2011: MALUS X ROBUSTA 'ERECTA', CULTIVAR OF MALUS, SHRUB/TREE, ROSE FAMILY, PROF. CHARLES SPRAGUE SARGENT, DIRECTOR, ARNOLD ARBORETUM, 125 THE ARBORWAY, JAMAICA PLAIN, MA, COLLECTED BY: SARGENT, C. S.

6857: SYRINGA KOMAROWII SSP. REFLEXA, SHRUB, OLIVE FAMILY, DR. ERNEST HENRY WILSON, KEEPER, ARNOLD ARBORETUM, THE ARBORWAY, JAMAICA PLAIN, MA, COLLECTED BY: E. H. WILSON



Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

2.10

Torreya grandis herbarium specimen. Courtesy of Harvard University.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

Alternative Histories

A radial version of the same timeline pushes the metaphor to embedded arboreal processes—the forming of rings in a tree (figure 2.11). But more important, new patterns are illuminated by the density gradient from center to periphery. The three eras of collecting become more prominent as the sparse accessions in early years are compressed into a smaller space. Subtle lines of accessions running through significant dates of the year are accentuated. They appear as concentrated rays within the circular geometry. The radial organization also suggests an entirely different kind of temporality: one that has an origin at some fixed point and then expands indefinitely into the future.

This radial image can be read against the linear one, which presents time as being infinite in two directions. The accessions depicted in linear form seem sparse in comparison. In the linear version, one can more clearly see increased collecting over the years, albeit with a narrowing in the 1940s. Moreover, practices seem to change dramatically across seasons in the second half of the twentieth century, transitioning from accessioning only in winter to year-round. Other patterns are less visible in the linear timeline. The dispersion of accessions in the early years makes it more difficult to note the intensity of wild collecting during the period of exploration and its symmetry with the period after the 1970s, when the arboretum began to collect externally again. At a more detailed level, a substantial gap in collecting on Christmas day appears clearly in the radial version, but disappears into the fringe of the linear image. This gap could be made more prominent by simply reordering the arrangement of months, but what other patterns would be shifted out of view?

Both the radial and linear versions obscure the exact number of accessions per day. A three-dimensional approach as demonstrated in figure 2.12 can help to make those more evident. Rather than being arranged solely by date, the 3-D image highlights every accessioned plant at the arboretum and exposes the rate of accumulation along a new z-axis. The resulting form is a cone. Moments of rapid growth in the collection appear as narrow sections, whereas periods of slower development flatten it out. While evocative in its shape, the 3-D visualization is more difficult to read. In fact, most of the patterns exposed by other visualizations are compromised in 3-D. Graphics overlap from opposite sides of the cone, the circumference of the yearly rings is visibly narrowed, and daily accessions are difficult to align by month and year.

The above examples of visualization are both interpretative and speculative. They present the arboretum as multiple. Each version of the place offers its own experience of the substantial collections brought together over a long history.

Histories Out of Place

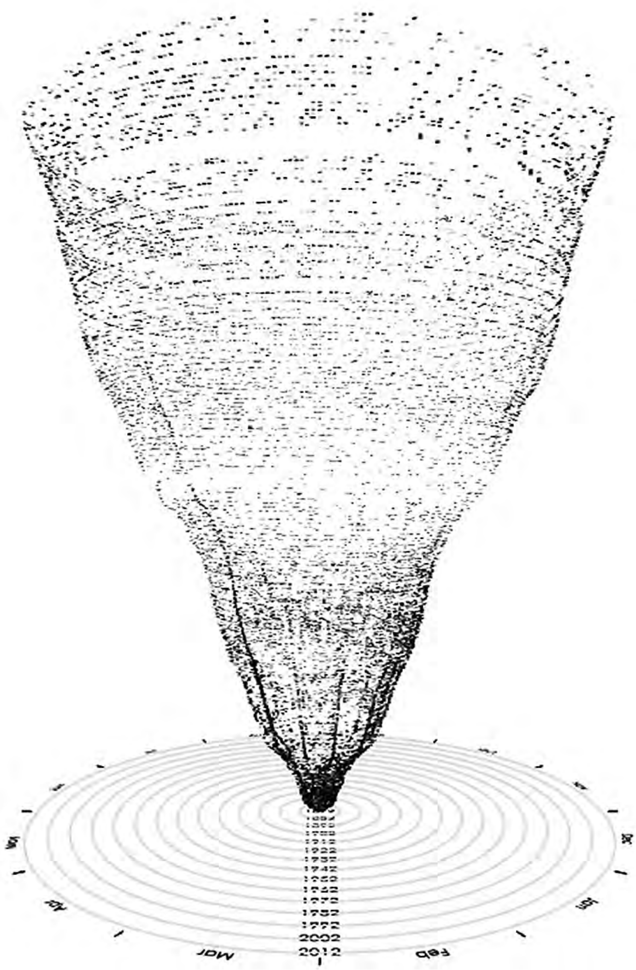
While the visualizations suggest different ways of making sense of the arboretum as a whole, they can also reveal telling details. Indeed, we can learn more about the kind of place that the arboretum is by inspecting components of the visualizations close up.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

2.11

Radial timeline of accessions to the Arnold Arboretum. The dots on the bottom edge represent plants with no accession date. Image by the author.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023



2.12

Three-dimensional timeline of arboretum accessions. Image by the author.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

In particular, it is useful to pay attention to apparent anomalies or glitches in the images. I call these "data artifacts."[29] In most work with data visualization, such irregularities are cleaned up. Like various kinds of data dirt, they appear to be simply out of place.[30] But data artifacts speak to the human history of their accumulation.

Consider, for instance, the rays of clustered accessions so prominent in the radial version of the timeline. A literal reading of these rays suggests that accessions arrived en mass on certain days, especially on the fifteenth of every month, the first of the year, and the first of July. Del Tredici, though, suggests that the rays are most likely technological artifacts. "If something came in (during) August 1942, I think BG-Base would output that [by] default as August 15."[31] Without a precisely recorded day of accession, BG-Base places accessions squarely in the middle of the month. The pattern is similar at the scale of the year; accessions appear unusually heavy on July 1, the beginning of the arboretum's fiscal calendar.

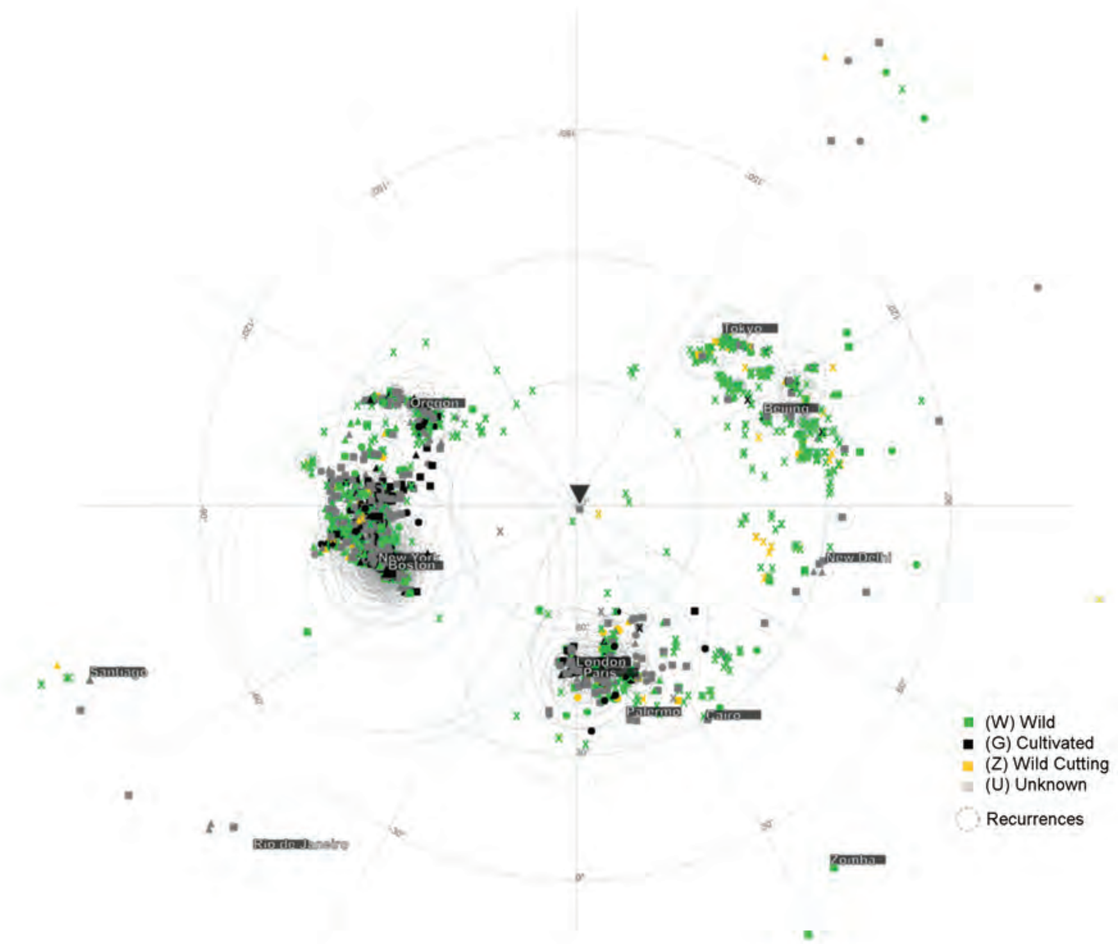
Deriving from various processes, such artifacts are often entangled with the contingencies of a place. Those mentioned above might be thought of as production artifacts, resulting from the technical conditions of data creation. Meanwhile, disciplinary artifacts might betray specialized ordering systems, and vernacular artifacts might be the result of dialects or local language uses. These various kinds of artifacts can be extraordinarily subtle and difficult to tease out, but visualization is an adept tool for bringing such conditions to the surface.

Data artifacts register not only local changes in technology, personnel, and organization but also broader cultural rhythms and events. Look closely and you can spot World War II as well as Christmas (mentioned previously) as gaps between denser periods of accessions. The first is manifest as a bald swath in the mid-1940s. The second is particularly noticeable in the radial timeline as a wedge of space radiating down the axis associated with December 25. Accessions from specific regions are affected by international relations too. Del Tredici recounts that "when [Richard] Nixon went to China, I started to get small little exchanges of seed packets and things like that."[32] Through data artifacts, we can see more than a collection of plants. Kyle Port, the arboretum's plant records manager, explains that artifacts betray the "personalities" behind the data.[33] Together, these personalities contribute to the aggregate sense of place generated through visualizations.

A Composite Place

One final visualization offers a view of the arboretum as a collection of places. In figure 2.13, the arboretum is presented as a set of locations extracted from the data. These are not the locations where plants were collected. Rather, they are the addresses of individual collectors. Mapping these data, done here in polar coordinates, results in an image of the arboretum's social network, with each dot representing the home or work address of a collector. The gray-dotted circles call attention to the areas with the largest

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023



2.13

Map of the addresses of arboretum collectors. Image by the author and Krystelle Denis.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

number of collectors, such as the arboretum itself. This way of visualizing the data suggests that the arboretum encompasses an extended landscape of collecting activity.[34] As in some of the previous visualizations, the colors represent provenance type, revealing the professional addresses of the collectors of wild and cultivated plants.

RETHINKING THE PLACE OF DATA

Today, popular media depict data as increasingly commonplace: ubiquitous tools for government, science, and business management.[35] Although there is a long history of scholarship on the *place* of information within discourses on cyberspace, cities, networking, interaction, and development, academic and popular discussions of data frequently downplay the significance of place.[36] Sometimes data and place are treated as incompatible concepts. Geographers Craig Dalton and Jim Thatcher argue that new practices with big data distract from attention to place. "Relying solely on 'big data' methods," they write, "can obscure concepts of place and place-making because places are necessarily situated and partial."[37] What if we learned to see data as situated and partial because of their place attachments?

Although place has been an important topic of interest in the social sciences, my readings of data are also influenced by cultural studies, particularly in the environmental humanities.[38] Lawrence Buell, a leading voice for ecocriticism, expounds on the multiple dimensions of place attachment in texts, including temporal and imagined conceptions of place.[39] My development of the notion of place attachment for data builds on these important precedents, yet it is grounded in readings of data manifest at the Arnold Arboretum.

Here I use *place* to mean an institutionally defined framework with social, technological, and spatial dimensions, in which data are created, displayed, and/or managed, and that reciprocally, is shaped by those practices. Indeed, data are not simply site-specific tools; they have the power to shape place. In common parlance, the term *data* can be used to mean secondary, digital representations of objects that hold scientific and cultural import. But data can also create an ontological "looping effect" whereby they help to shape the practices and institutions that create them.[40]

Do place attachments still hold at the scale of big data? Accessions data at the Arnold Arboretum certainly don't conform to present-day definitions of *big* as high magnitude in a variety of dimensions.[41] Instead, researchers and other staff at the arboretum demonstrate the kind of close relationships with data that contemporary big data approaches were meant to replace, since practices like those carried out every day by Arnold Arboretum botanists require time and proximity that are too often dismissed as expensive and unnecessary.

Having said that, if we consider big data as an epistemological and performative shift in ways of doing research, with a long history involving data sets that were previously unmanageable, we might say that the Arnold Arboretum has been edging toward

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

big data for over a century. In the nineteenth and early twentieth centuries, arboretums—like libraries, museums, and zoos—held the big data of their day. Institutions like the Arnold Arboretum prefigured big data by drawing together representative specimens from far and wide. The most ambitious of these institutions sought to establish themselves as comprehensive models of the world.[42] As with contemporary holders of big data, such institutions of collecting continually outstripped strategies for managing all the records necessary to organize, preserve, and study their subjects. The arboretum's many successive eras of data collection illustrate, better than most, a variety of place attachments in data, regardless of their magnitude.

CONCLUSION

We should learn to see data as cultural forms that are situated socially and technologically, but also in place. Although data are reliably transferred across global communication networks everywhere, they remain marked by local artifacts: traces of the conditions and values that are particular to their origins. Accepting this claim necessitates a significant shift in our expectations of digital data given that the digital was invented to be independent of any substrate.[43] In fact, all data—not just those created at arboretums and other sites for documenting nature—can be read distinctively through their attachments to place.

Each of the place attachments explored in this chapter suggest a different way in which data can be local: by being about, in, from, or even generative of place. Taken together, these four ways of probing data offer a model for how to read data from a local perspective. But this should not be mistaken as a formula for engaging with data anywhere. My methods were developed in situ, with the particular place attachments of the arboretum at hand. Similarly, I encourage readers to challenge these and other settled conceptions of the relationship between data and place by making place a part of their data analysis as well as data presentation.

Exploring the possible relationships between data and place can help us question the wisdom of centralized models of data management. As this chapter has shown, thinking about data as mobile, immutable, and generally detached from place can obscure important ways in which data practices rely on local knowledge as well as experience for meaningful interpretation and responsible use. When taken out of place, data can come to be seen as either the view from nowhere or nothing more than data dirt.

In the next chapter, I explain what happens when data from many places are brought together in the form of data infrastructures. Using the example of the DPLA, a composite collection of digitized media from libraries, museums, and archives across the United States, I illustrate some of the ways in which even displaced and agglomerated data retain traces of their origins, embedded in classifications, schemata, constraints, errors, absences, and rituals that resist simple translation or normalization.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023

3

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2014835/9780262352215_cac.pdf by AUSTRALIAN NATIONAL UNIVERSITY user on 08 March 2023