A Design Perspective on Data

Melanie Feinberg
School of Library and Information Science (SILS)
The University of North Carolina at Chapel Hill
mfeinber@unc.edu

ABSTRACT
Empirical studies invariably show that data generation is situationally contingent and interpretively flexible, even when data is collected automatically. This essay situates data generation within a design perspective, demonstrating how data creation can be understood as a multilayered set of interlocking design activities. By showing how data is infused with design, this paper argues that any "use" of data represents a continuation of its design. We are always designers of data, never its mere appropriators.

Author Keywords
Data; design; infrastructure; metadata; materiality

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

INTRODUCTION
It is well understood that data involves more than mere recording of facts. Researchers across disciplines have described the mixture of human motivations, historical traditions, environmental and technological affordances, and particular goals in which data is conceptualized, structured, captured, aggregated, and analyzed [see, for example, 3, 12, 17, 18, 25]. In HCI, user studies describe the richly situated environments in which data is constituted [as in 2, 16, 21, 26]. Most recently, studies of personal informatics detail how data collected by automated mechanisms (as with fitness trackers or household sensors) is dependent on context for its meaning, in accordance with user behavior that might be systematic and consistent, or that might equally be fluid and improvisational—or sometimes both [for example, see 5, 6, 8, 23, 27]. For instance, a step counter on a smartphone might be implemented by one person as a meticulous, comprehensive log, while another person uses it an opportunistic and haphazard recording of partial activity. Although the counter is just recording steps, its deployment is situational, creative, and flexible. Design of data doesn't stop with the design of a step-counting feature; it proceeds through the data collection patterns of smartphone users, and it continues as counts are generated and aggregated over time.

CHI 2017, May 06-11, 2017, Denver, CO, USA
© 2017 ACM. ISBN 978-1-4503-4655-9/17/05…$15.00
DOI: http://dx.doi.org/10.1145/3025453.3025837

What does it mean, then, to "use" step count data as an element of some subsequent design, perhaps to visualize activity patterns in a neighborhood or to suggest health improvement goals? In this paper, I illustrate how all "use" of data involves its design: how data is never merely the input to design, but always its object. I show how data creation can be understood as a multilayered set of interlocking design activities, and how data use is the continuation of these activities.

While "data" may appear conceptually simple—a data point is just a value, often a number, like the number of steps taken each day—any data requires some kind of *conceptual infrastructure* to have meaning. The number of steps involves the designation of a form of measurement (steps) and a scale of measurement (positive integers of whole steps). Data also involves *collection processes* that individually implement the conceptual infrastructure (how each step is counted) and *aggregation processes* that integrate many acts of independent collection (the association of each step with a particular device across established time units). To make my argument, I trace the role of design in these areas: infrastructure development, collection, and aggregation. I demonstrate how design figures into these activities using a variety of examples. (Although some examples are drawn from a research study, this paper is not a report of findings. It's an essay in which research findings are deployed as one form of evidence.) To provide context for my discussion, the next section briefly reviews accounts of data collection as an interpretively flexible, locally contingent performance, as presented in HCI and aligned fields.

RELATED WORK
Data collection has become a pervasive, everyday activity. Applications routinely collect data to algorithmically shape interaction experience, and individuals increasingly elect to collect data on themselves to inform their own behavior. Accordingly, descriptive accounts of data creation and use have proliferated in the CHI community. Many studies focus on personal tracking data, and the moniker "human-data interaction" has been proposed to understand how people interact with their personal data [5]. When researchers look at practices of personal data collection and use, they observe a diverse set of creative, flexible approaches to data creation, tightly integrated within a local context [6, 8, 27]. Rooksby and colleagues, for example, describe how people employ fitness tracking devices to

collect data in highly individual ways, even though the actual data is collected automatically by sensors. Someone might choose to wear a fitness tracker only for certain activities or at all times, in effect creating different kinds of datasets to align with their own, dynamic needs, preferences, and understanding [23]. Similarly, Tolmie and colleagues describe the deployment of household sensors in three households and the variations in behavior that give rise to variations in data [27]. Leaving the garage door open, which makes the kitchen temperature drop: this is a data creation event, just like deciding when to wear a fitness tracker. Others have made similar observations in the context of government data [2, 21]. McMillan and colleagues, for example, interview stakeholders in northern European cities to understand how "the production and use of data is shaped by local contingencies of influence, power, money, and bureaucracy" [21].

The creative, situated, and interpretive nature of data collection is well established across disciplines. Goodwin describes how archeologists apply the Munsell color chart to earth samples [12]. Latour recounts how botanists distill the forest into a cabinet of specimens [18]. Bowker reveals differences in collection and testing of lake water samples over time [3]. Star and Griesemer note competing concerns of trappers and naturalists collecting biodiversity specimens in the Sierra Nevada [25]. All these studies and many others demonstrate how the practice of data collection—even in science—is informed by disciplinary convention, situational contingencies, and individual flair. Notably, interpretive flexibility in data creation occurs despite the use of standardized schemas, vocabularies, and protocols to enforce consistent, semantically interoperable data. With a professional tradition dating from the nineteenth century, libraries have developed an incredible array of guidelines, processes, and roles devoted to reliable data collection. Nonetheless, Carlyle notes how complex, dynamic realities constantly challenge catalogers to interpret rules anew—even for matters that seem banal and straightforward, like determining the author of a book [4]. Kansa, Kansa, and Arbuckle describe how zooarcheologists working in the same region in Turkey, using the same system to describe tooth wear on fossils, nonetheless structured tooth data so differently that skilled human editors could not easily aggregate it [17]. Jackson and Barbrow discuss the insufficiency of an extremely detailed collection protocol for a single ecology lab to enable frictionless data collection [16]. From a "big rock in the middle of the stream" to bears "disrupting the collection equipment" the world presents an unruly environment that cannot be circumscribed by standards but must be interpreted in light of them. Across such accounts, the same conclusion arises: although data collection might be mundane and tedious, it is also dynamic, creative, and unpredictably diverse, even when data is collected without apparent human intervention.

Given this interdisciplinary consensus, it is significant that work practice often contradicts it. For example, standard practices of information professionals (people who generate, curate, and aggregate data as a primary job responsibility, including librarians, archivists, and data managers) continue to approach interpretive flexibility in data creation as a problem to be solved. As noted in the preceding paragraph, scholars of information studies have long acknowledged the role of interpretive judgment in data creation. For example, in 1968, Wilson argued that a document's subject (its aboutness) could not be conclusively determined [30]. As Furner has observed, there is no scholarly opposition to Wilson's argument [9]. Nonetheless, professional practice continues to operate as if documents "have" subjects that can be accurately and consistently identified. What causes this conceptual dissonance? Professional goals and values align poorly with empirical realities of data variability and interpretive flexibility. In information professions, data utility is seen as dependent upon globally consistent meaning [13, 15, 32]. Accordingly, like Sisyphyus eternally pushing that rock up the hill, data professionals strive to approximate universality and objectivity in data collection, even if such goals can never be achieved. In this view, if creative interpretation can't be eliminated in data collection, practice rules and guidelines might at least minimize it.

This inconsistency—where one knows conceptually that data is interpretively fluid and yet acts as if it were, might be, or should be objective and universal—is not limited to information professionals. In the studies of archeologists, botanists, ecologists, and so on cited earlier in this section, scientists, also, cling to an ideal of globally reliable, semantically interoperable data collection. Why? Their professional goals and values, also, align poorly with empirical realities of interpretive flexibility in data creation. Just like information professionals, scientists attempt to account for and ameliorate situational conditions that complicate data interoperability. They make these efforts so that, after data is collected, they can aggregate and "use" (or reuse) it, without considering how the character of the data might continue to evolve through its aggregation and use.

This conceptual dissonance can also surface in HCI. Descriptive studies of data creation in HCI emphasize its creative situatedness, and design research has embraced interpretively flexible design outcomes [see, for example, 10, 24, 29]. But design projects in HCI can omit the work performed on data, making it seem as if data were a stable material to be "used." For example, Gaver and colleagues built beautiful, custom "datacatcher" devices to convey socioeconomic data about a user's location [11]. Their paper focuses on the effort to design and build the datacatchers, to distribute them, and to document their use. The team's activities with the data conveyed through the device are minimally documented. The project's data work is briefly described as pulling from "hundreds of data sets from 14 online sources," with "templates to transform numerical or category data into sentences." Torres, O'Leary, and Paulos created evocative art objects to display data sources, such as a set of fans whose speeds varied depending on Twitter mentions of candidates for president of the United States. Their paper also uses a minimal approach to describe their work with data. For example, the paper does not elaborate upon the design choice to employ “mentions” as the data element (and not, say, the content of the mention or the sentiment of the mention) [28]. Worthy, Matthews, and Viller deployed a technology probe to spark reflections on data gathering and the Internet of Things [31]. The probe collected sensor data from participant households, but the paper focuses on the general idea of data collection rather than what was actually collected.

In the following sections of this paper, I argue for a view of data as a multilayered design artifact, which implies that we can never incorporate data into a product without redesigning the data itself. To understand data from this perspective, I draw on ideas of thing-design, use-design, and design after design from Johan Redström [22]. Thing-design focuses design on an object: a chair. Use-design focuses design on an activity: sitting. Design after design involves the “user” of an artifact adapting it to a new purpose. By installing, configuring, and adapting software, the “user” of a computer is also a “designer” of her personal experience with it. Redström describes design after design as a series of processes in which the product of one design activity becomes the material for subsequent design activities: a textile is designed and then used as material for a dress, for instance [22]. The A. Telier group draws on these ideas to propose the idea of infrastructuring a goal of participatory design. Infrastructuring involves creating conditions under which dialogues to imagine new design possibilities are facilitated [1]. The A. Telier group focuses on infrastructuring as enabling dialogue between people, drawing on an older sense of the word *thing* as a meeting place to resolve disputes. But such conversations can occur between people and design materials as well. I use these concepts to present data as a dynamic cascade of design decisions, with each “product” serving as “material” for the next move in an ongoing design chain.

THE ROLE OF DESIGN IN DATA INFRASTRUCTURE

All data relies on decisions regarding what to record, how those values should be collected, and the form in which those values should be expressed. These decisions constitute the infrastructure under which data is initially created, subsequently aggregated, and ultimately used. For instance, weather is typically represented with temperature data. We all know that temperature is measured with a certain scale (usually Fahrenheit or Celsius). But the collection of temperature data relies on other decisions as well: the device used for measurement (such as a thermometer), the placement of that device (on the ground, in the air, in the sun, in the shade), the times at which measurements are taken. Moreover, the selection of temperature as a useful property to understand weather is itself part of the data infrastructure.

In the vocabulary of relational databases, data infrastructure involves the selection and definition of entities, attributes, and value parameters, plus the processes established to facilitate data creation. For the example of temperature, the entity might be weather, with an attribute of temperature, as expressed in degrees Celsius. Although a typical entity-relationship diagram wouldn’t specify data collection procedures, such processes are often defined and documented as part of data generation protocols. For example, the protocol for ecological field data described by Jackson and Barbrow details not just what kinds of data to collect and the form in which observations should be collected but the processes for doing so [16].

Many elements of data infrastructure might also be described as “metadata.” I’ve chosen not to use “metadata” because that term can be employed to indicate a wide range of data-related notions that are both more and less than the idea of data infrastructure defined here. One might refer to the date that a photo is taken as metadata in both a general sense (the date of capture as a kind of data to describe photos, just as the temperature is a kind of data to describe weather) and in a specific one (the actual date and time at which a specific photo is taken). “Date taken” is metadata about photos in general, and “July 10, 2016” is metadata about a specific photo. In the context of data infrastructure, I mean the general sense and not the specific one.

In HCI, the development of data infrastructure is not commonly considered as a design activity, perhaps because such decisions often seem banal and obvious, or because they seem more associated with subject-matter expertise and scientific methods than with a design perspective. The use of temperature to understand weather, and the means in which temperature is measured, seem a matter for meteorologists to specify, not for designers to interrogate or imagine differently. I’ll use a set of examples of data infrastructure from a specific context—online dating services—to show how design is relevant to these kinds of decisions, no matter how mundane or scientific they appear.

Many online dating services support the creation of user profiles to facilitate both retrieval (finding a set of potential dates) and selection (determining the best option in the set of matches). Users are asked to describe themselves along a wide variety of attributes, many of which are similar across services. Two of these common attributes are body type and ethnicity, physical characteristics that might—naively—seem simple and straightforward, not a matter for design. These attributes are typically expressed via a set of controlled values, which vary across dating sites. Tables 1 and 2 show controlled values available for daters to describe themselves according to these attributes in five dating services. I’ve used these online dating examples many times to explore the space of data infrastructure in the classroom. Students’ responses are always the same. When

Match	OK Cupid	Black Planet Love	JDate		Gluten-Free Singles
No answer	Rather not say	Athletic/muscular	Lean/slender	Proportional	Slim
Slender	Thin	Disabled	Average/medium build|	Ripped	About average
Athletic and toned	Overweight	Large frame	Athletic/fit	Rubenesque	Athletic
A few extra pounds	Average build	Medium	A few extra pounds	Small frame	Pumped up
About average	Fit	Toned	Large/broad build	Soft	Have some extra pounds
Heavyset	Jacked		Cuddly	Stocky	Big and lovely
Stocky	A little extra		Firm and toned	Voluptuous	
	Full figured		Husky	Zaftig	
	Curvy		Petite	Muscular	
	Used up		Portly	Modelesque	

Table 1. Body Type options in five online dating sites, as collected in July, 2016. Match and OK Cupid are large sites that serve a general population, while Black Planet Love, JDate, and Gluten-Free Singles target particular characteristics of race (Black people), religion (Jewish people), or lifestyle (people who follow a gluten-free diet).

Match	OK Cupid	Black Planet Love		JDate	Gluten-Free Singles	
Asian	Asian	Asian/Pacific Islander	*Note:* A separate item for Additional Ancestry lists about 15 cultural groups, primarily of African, Caribbean, and Latin American origins.	*none*	African	Indian
Black/African descent	Middle Eastern	Black/African American			African/American	Latino
East Indian	Black	Hispanic/Latino			Asian	Mediterranean
Latino/Hispanic	Native American	White			White/Caucasian	Middle Eastern
Middle Eastern	Hispanic/Latin	Other			East Indian	Mixed
Native American	Pacific Islander				Hispanic	
Pacific Islander	Indian					
White/Caucasian	White					
Other	Other					

Table 2. Ethnicity options in the five online dating sites, as collected in July, 2016 (Black Planet Love labels this as Race, not Ethnicity). These are implemented as checkboxes except for Gluten-Free Singles, which uses a menu.

I ask them to think about characteristics like body type and ethnicity in a general sense, these properties seem like relatively stable concepts that can be defined in a general, neutral, and objective way. The project of creating data infrastructure for these characteristics, such as lists of controlled values, seems like a scientific kind of project, not a design one, and it doesn't seem problematic to think about creating a universal set of such values that would work across, say, all online dating services. Then I distribute the controlled vocabularies that are actually used to describe body types in these dating services (see Table 1). After students stop laughing at values like Big and Lovely, Jacked, and Cuddly, I ask if the variation between dating services is problematic. Shouldn't there be a single authoritative means of capturing this information for all situations? After seeing these real implementations, students always respond that the variation is not problematic at all: this diversity represents different understandings of body types for the different communities represented by the different dating services. Moreover, the idea of body type itself is revealed, as implemented in dating sites, to be different than the students might have initially conceived. Body type is not neutral, objective, and general but personal, performative, and situational: it has to do with how you feel about your body and how you want others to perceive your attitude toward your physique, rather than your actual proportions. It is easy for students to imagine selecting different values for body type in different dating services—to be Slender in Match, Average in OK Cupid, and Proportional in JDate, for example—because the infrastructure is different, but also because the infrastructure indicates that the community and accompanying situation is different—and so daters, too, will have different identities in each.

To continue this discussion in the classroom, I observe that it is of course possible to represent the concept of body type in other ways, perhaps as a set of quantitative measurements for height, weight, and the circumference of various body parts. Wouldn't that have some advantages? The data would be less subjective and more interoperable [13, 32]. Invariably this is derided as a terrible idea. Someone often comments that people would lie. Imagine, I say, a system where your bathroom scale updates your profile automatically, along with sensors in your clothes to continually measure your waist, hips, and other body parts (not far-fetched with the Internet of Things). People look aghast. That would be even more terrible! That's weird, I observe. You mean you don't want accurate data? It's not that, the students say. What they mean is this: they've realized that the idea of body type as a performative, personal concept is more useful and illuminating in the dating context than body type as a set of quantitative measurements. Accuracy is both relative and insufficient.

The ultimate lesson from this example, though, is not that certain kinds of data infrastructure (the performative, situational idea of body type) are designed, while other implementations (the quantitative, universal idea of body type) are not designed, but that *all* data infrastructure is designed. As noted earlier, Redström uses the example of a chair to illustrate the distinctions between "thing-design" and "use-design" [22]. Although designing a chair involves specifying a thing, it also involves the suggestion of certain activities to be performed with that thing: ways of sitting. The process of design might emphasize characteristics of the artifact or its uses. Determining the infrastructure for quantitative data collection generated by conventional-seeming automatic means (like weight from a scale) might appear to involve few design decisions on the artifact axis, but such decisions of data infrastructure nonetheless specify quite distinct modes of use. Height and weight can be collected in a manner that is perhaps less subject to human interpretation than selecting from one of the Body Type vocabularies. But the *design choice* of using height and weight is not disinterested or neutral. Using height and weight to represent body type circumscribes a particular, situated idea of what "body type" means, and the activities that can be performed with height and weight data are different than the activities that can be performed with selections from the Body Type vocabularies. Deciding what data to collect, the form in which it is collected, and the means of collection—these are design decisions, in that they entail particular activities with the resulting artifact.

The data infrastructure for ethnicity in the five online dating services illustrates this further (see Table 2). Particularly in the United States, the primary market for these dating sites, ethnicity and race are politically and socially fraught concepts. The experience of race and ethnicity varies widely in the U.S. population. White people don't tend to notice social privileges accorded to their race, but people in historically oppressed groups are constantly reminded of racial disparities. To specify race/ethnicity categories in dating profiles and enable potential matches to be excluded or included based on these characteristics is to acknowledge these realities and to express a position on the social salience of particular group identities. The differences between the five dating services demonstrate this. Black Planet labels this category Race, and not Ethnicity, and the values used for the Race category position physical appearance as primarily determining the social experience of race. Cultural experience is secondary to perception of physical characteristics in Black Planet—it's how other people see you that is most important. In contrast, Match, OK Cupid, and Gluten-Free Singles distinguish groups according to cultural distinctions as well as physical appearance: Some East Indians (Indians, Pakistanis, Bangadeshis) might appear similar in physical appearance to some Middle Eastern people (Arabs, Egyptians, Iranians), but these groups are separated as culturally distinct. In this perspective it is not surprising that JDate, which targets a single cultural group—Jews—lacks a category for ethnicity or race, although there are non-white Jews of different national heritages. To include such a category would be to deny the primacy of Jewishness as a group differentiator. As with body types, to quantify such data is to miss the point: representing race by, say, DNA markers of ancestry would be to express *an entirely different concept,* one that facilitates quite different activities. Choosing to collect "data" in a particular form, with a particular process, expressed in a particular manner: these are design decisions, no matter how mundane and conventional such decisions might appear.

If we take the A. Telier's group's perspective of design as infrastructuring to facilitate dialogue-things, the conversations possible with race-as-DNA data are different than the conversations possible with data from ethnicity-as-user-selected-labels [1]. The conversations facilitated by data infrastructure are not limited to users of the data created with that infrastructure: to people trying to find matches on dating sites. Other conversations involve the use of data infrastructure to collect data (in this example, people filling out profiles) and to aggregate data (the construction of sets of profiles within and across dating sites.) Data infrastructure provides a set of conditions under which design after design can occur. The next sections look at collection and aggregation from this perspective.

THE ROLE OF DESIGN IN DATA COLLECTION

In this section, I describe a view of data collection as a form of design after design, where each act of data creation involves a dialogue with the data infrastructure [22]. Redström's concept of design after design aligns with empirical accounts of data collection. As discussed in the related work section, people collecting data interpret data infrastructure creatively, flexibly, and situationally. Data infrastructure does not *determine* data; it provides conditions under which people *create* data. In the context of data collection, data infrastructure is a design material.

As a material, data infrastructure is manipulated in diverse ways by different actors. An online dater might select a body type aspirationally, realistically, or pessimistically. Relative consistency of interpretation within a particular community can also change over time. Library catalogers today employ greater specificity and exhaustivity of indexing when applying subject headings to documents than in earlier eras; there are more headings, and headings are more specific, than in the past [7].

Nonetheless, it remains slightly strange to conceive of data collection as design, just as it remains non-obvious to think of customizing one's computer as design. One might recognize *in the abstract* that customizing one's computer is a form of design, but it doesn't much seem like it when you're installing a new application or changing settings. The same holds true for data collection. Although data collection does involve creative decision making, it is simultaneously banal and tedious. Think about your personal data collection: perhaps you track your menstrual cycle or take regular blood pressure readings. Without thinking much about them, you probably see these activities as rote and mechanistic. But if you stop and consider the details of your practice, you're also performing these tasks with a degree of creativity and flexibility. I, for one, know that I interpret the data labels in the Clue menstrual tracking app in a way that is unique to me. I *am* manipulating data infrastructure as a material in a form of design after design. And yet—I wouldn't call it design when I'm doing it.

A design perspective on data collection acknowledges empirical realities of practice and enables innovative reconceptualizations of data creation and use. To continue this argument, I use evidence from a dataset created as a semester project in two separate courses for graduate students in a master's program in information studies, conducted in spring and fall of 2015 (Course A and Course B). The dataset comprises records that describe a particularly complex information object: video games, which encompass tremendous diversity in form, structure, and content. The data infrastructure for the study included a schema, documentation, and vocabularies developed as a set of standards for video game description [19, 20]. Using these data standards, student participants each described 7 or 8 video games of their choice and 3 games common to all participants. After creating the data, students analyzed the data for their class and composed 3000-word essays that assessed interpretive flexibility in the aggregated data.

Combined, the dataset from the two courses includes 282 records, created by 26 student participants (15 in Course A and 11 in Course B). The schema used to describe the video games includes 45 elements for describing each game, including such disparate attributes as Mood, Platform, Price, Digital Rights Management, Visual Style, Networked Features, and Publisher. The dataset and its accompanying essays are useful evidence because the full provenance of each record is known: the data creator, the data infrastructure employed, and the immediate context (each course). Moreover, while some schema elements appear clearly interpretive (such as Mood or Visual Style), others appear much less so (such as Platform and Price). Like one's use of a menstrual tracking app or blood pressure monitor, recording a video game's price seems banal. Accordingly, Price is an illustrative example.

In the video game schema documentation, instructions for the Price element are:

Definition: The manufacturer's suggested retail price (MSRP) at time of initial release in the region where the game was released.

Instruction: Determine the Manufacturer's Suggested Retail Price (MSRP) from the CSI. Record the price with the currency, source, and the date when this information was acquired. If unknown, specify "Unknown".

Example: 59.99 (USD, Amazon.com, 2014-03-25)

"CSI" means "chief source of information." The schema documentation lists primary and secondary choices for the CSI, including the game box (for games sold in separate packaging), the site where a downloadable game was obtained, and secondary game sites (including professional media, fan sites, and Wikipedia).

For one of the 3 common games, Final Fantasy 7, the following was recorded for the Price element in Course A:

- Blank
- Unknown (2 times)
- 4.99
- 11.99 (3 times)
- 14.99 (6 times)
- 19.00
- 149.99

For the 12 entries that included price information, 7 followed the basic structure of the suggested price format. Of the 5 remaining entries, 2 omitted a source and date, one omitted a source, one omitted a date, and one used a full URL instead of the source name. Of the 10 entries that included a date, 8 indicated when the price information was acquired (in spring of 2015), as the schema documentation directs. One of the others was the release date. The final date was not in 2015, so it wasn't when the price information was acquired, but it wasn't the release date, either—its significance is unclear.

The following price information was recorded for Final Fantasy 7 in Course B:

- Blank (5 times)
- Unknown
- 11.99
- 14.42
- 14.99 (3 times)

Of the 6 entries that included price information, none of the entries used the recommended price format. Three used a full URL instead of the source name. Three didn't include a source. Five entries included a date; all of these indicated when the price information was acquired (in fall of 2015).

Where is the design in recording the price of Final Fantasy 7 as Unknown, 11.99, 14.99, blank, or something else? This seems like a fact that can only be accurate or inaccurate, but I suggest that it is something else: part of a dialogue with the data infrastructure regarding what games and prices are (thing-design) and what activities might be facilitated in describing price (use-design).

In the video game schema, a descriptive element like Price implies a certain level of abstraction in terms of documenting a game: an edition for a particular platform (such as Playstation 4 or iOS). The definition of Price in the schema documentation clarifies that price applies to a "release" in a "region" (the schema includes additional elements for Retail Release Date and Region Code). A release in a region suggests a level of abstraction similar to that of "manifestation" in the Functional Requirements for Bibliographic Records (FRBR) model [15, 19]. A FRBR manifestation is a set of objects that include the same intellectual content and physical characteristics, such as the set of original LPs for Marvin Gaye's *What's Going On?* on Tamia Records. For video games, it is reasonable to see a manifestation as something like Final Fantasy 7 for the Playstation 1 as released in the United States.

But the schema documentation doesn't *specify* a level of abstraction for creating data, and other elements seem to apply equally to all editions, or to all editions with the same "intellectual content" but perhaps different "physical characteristics" (the "work" or "expression" level in FRBR). Mood and Visual Style, for example, would presumably be the same for versions of Final Fantasy 7 on Playstation or PC, and also the same for versions in Japanese or English. There is room for a data creator to make different choices about level of abstraction.

Price intersects with level of abstraction more specifically as well. The schema directs creators to record the MSRP at the time of initial release. The first edition of Final Fantasy 7 for Playstation 1 was released in 1997, but other editions of Final Fantasy 7 were available in 2015, when the data was collected. In developing the schema, Lee and colleagues sought to facilitate the information-seeking needs of a wide variety of user types: serious and casual gamers, parents of young players, developers and other industry professionals, collectors, scholars and educators, and curators/librarians [19]. In evaluating the schema, Lee and colleagues surveyed over 1200 respondents about the usefulness of each element in the schema, and price information was determined to be useful by the largest percentage of people (82 percent) [20]. But which price is useful for whom? The price that a particular seller was asking to purchase a particular copy of the game when the data was created? The price that the manufacturer suggested for a particular version at the time of the game's release—which no one may have paid? Collectors and scholars are interested in games for different reasons than gamers and parents, and different kinds of prices may align better or worse with different uses of the data. More importantly, different data creators might come to different, equally defensible decisions based on the same goals and evidence.

The diversity of approaches that Course A and Course B data creators took to implement price information demonstrates this. Three creators described multiple editions in their records; these creators left Price blank or used Unknown, enabling them to avoid listing multiple prices. Creators that indicated a price of 14.99 documented the first U.S. release of the game, for Playstation 1. Creators that indicated a price of 11.99 documented a PC version of the game available for streaming in 2015, when data was collected. No one described the first release for PCs, from 1998 (with a price of 77.98). The strangely high price (149.99) is associated with the first U.S. release and is compatible with purchasing the original discs for that release in 2015.

Individual data creators tended to apply a reasonably consistent interpretation of price throughout their own records, which indicates that these decisions were principled, not accidental. For example, data creator A02 (from Course A) used the format "price-currency-release date" ("4.99 USD 1997-01-31") for all entries. This structure aligns with an interpretation that "MSRP" implies price information is necessarily associated with the release date, no matter when that information was recorded (in other words, the manufacturer sets the MSRP at time of release, and it doesn't change after that).

For apparently objective elements like Price, creators were unsettled by the diversity of approaches to implementing the data infrastructure of the schema, even if they were individually consistent and principled in their actions. In their essays, many suggested changes to the schema and its documentation to constrain the interpretive flexibility they observed. Creator A09 proposed that that "the more objective the element, the less semantic diversity [should be] allowed," and others made similar assessments. Many creators suggested that the schema documentation should clearly specify a level of abstraction for description.

Such proposals constitute a typical response to situations like the Price element: to build more data infrastructure to make the data collection task require less interpretation and restrict opportunities to adapt the intended thing-design of infrastructure developers [an approach advocated in 17]. Indeed, a subsequent version of the video game schema specifies a set of entities (such as series, franchise, and edition) to establish and relate video game versions. But increasing the complexity of data infrastructure makes it more difficult for creators to implement, thus displacing the locus of interpretive flexibility without diminishing it. It is also possible to constrain the intended use-design of the data: if we were only selling games, then the price is the seller's current offer, about which there is less variability. This limits the data's utility for aggregation and reuse in other contexts, however—one of our primary goals with data. Moreover, experience demonstrates that variability will never be entirely contained. Any position that data creation can be perfectly specified contradicts our vast, cross-disciplinary empirical evidence regarding data collection practices.

In contrast, understanding data creation as a process of design after design allows us to imagine working *with,* rather than against, inevitabilities of interpretive flexibility. For example, instead of designing data infrastructure that specifies a correct level of abstraction for a certain kind of data (to establish that prices *must* refer to individual copies of a game or that moods *must* refer to all versions of a game release), one could design data infrastructure that enables data creators to set the level of abstraction for any statement (so that a price *could* refer to a large set of versions and a mood *could* refer to one person's individual copy). In other words, we might purposefully design data infrastructure to function more directly as design material—to support a range of possibilities for data creation, just like we design computer interfaces to function as material for new ways of working and living with devices.

This kind of approach might seem strange in regards to data—isn't data useful when it is reliable and consistent, not when it is interpretively creative and potentially ambiguous or contradictory? If we aren't sure what data means, then how can we trust it, or even use it, when it's aggregated? The following section demonstrates how aggregated data can be interpretively creative and still useful, if we think about utility in a slightly different way.

THE ROLE OF DESIGN IN DATA AGGREGATION

The "data" we use is invariably the product of multiple acts of data creation. Just as all data requires infrastructure to be intelligible, it requires aggregation to be meaningful. Tracking the number of steps you take in a day is worthless unless you compare that number to a reference standard, and tracking the steps taken in one single day provides little information about overall activity or fitness. Although all data involves some type of aggregation, the digital environment increases the potential scale of aggregation tremendously. (A common assertion regarding "big data" is that this increase in scale is so vast as to be revolutionary.)

The standard approach to aggregation involves making each data source compatible in both syntax and semantics [13, 32]. One data source is then mapped to another, or multiple data sources are mapped to a common structure. Standardized data infrastructure and controlled implementation of data collection are often employed to facilitate aggregation. Especially in large-scale aggregation, datasets might be "cleaned" to transform inconsistencies or remove problematic variation. Cleaning can be automatic or manual. Kansa, Kansa, and Arbuckle describe the extensive work of human data editors, with the assistance of some automatic tools, to aggregate 17 zooarcheology datasets from central and western Neolithic Turkey [17]. Although data creators had employed similar data infrastructure, they had implemented it differently, and mapping from one dataset to another was not trivial.

The example of Final Fantasy 7's price from the previous section illustrates this well. Data creators made different decisions in implementing the data infrastructure that resulted in different understandings of price and the entity it refers to (a unique item; the game as released for the Playstation 1 in the U.S. in 1997; the game as released for the Playstation 1 in the U.S. 1997 as available in 2015 for streaming on a computer; and so on). Some data creators rejected the data infrastructure's conceptualization of price, leaving the information blank or using the value Unknown.

To clean this data, a typical approach would take the predominant level of abstraction for video games associated with price and map all the data to that level of abstraction: for example, to make price associated with a set of versions released on the same date in the same locality for the same platform (the version for Playstation 1 in the U.S. in 1997 or the streaming version for PCs in all localities in 2015). Records that included data for multiple such versions could be split apart. Records that included data for individual items could be mapped upwards to the appropriate level of abstraction. Such mapping operations enable data to be more reliably compared and are often described as reducing noise or errors. But these cleaning transformations necessarily involve simplifying assumptions. In this approach, some data would be thrown away, such as price information for individual items (the 149.00 price for a current set of individual Playstation 1 discs).

Understanding aggregation as yet another layer of design in data creation, I suggest, enables us to imagine alternate approaches to data integration and use. One kind of alternate approach might consider how to take advantage of variation in aggregated data, instead of suppressing variation. Although ambiguity in the level of abstraction associated with price makes it difficult to perform some kinds of operations on the video game data, such as tracking price changes over time, this same ambiguity provides useful, interesting evidence of another sort: it empirically shows how video games such as Final Fantasy 7 are complex digital entities with many versions that differ in complex, intersecting ways. The ways that data creators have implemented price provides evidence for understanding the kinds of characteristics (current availability, historical importance) that mark particularly salient version sets.

Another example from the video game data illustrates this differently. This example looks at the Mood element, which is described as follows in the schema documentation:

Definition: The pervading atmosphere or tone of the video game which evokes or recalls a certain emotion or state of mind.

Instruction: Identify the prevailing mood(s) of the game according to the CSI; for most games, the experience of playing the game or watching the gameplay video may be the most reliable source of this information. Select the most appropriate term(s) from the CV for this element. If no mood is applicable, write N/A.

"CV" indicates a controlled vocabulary, which was created as a part of the video game data infrastructure. Terms in this vocabulary included:

Adventurous	*Horror*	*Humorous*	*Sad*
Aggressive	*Humorous*	*Mysterious*	*Sarcastic*
Competitive	*Imaginative*	*Peaceful*	*Sensual*
Comradery	*Immersive*	*Quirky*	*Solitary*
Cute	*Intense*	*Romantic*	

Here is how data creators in Course A applied these terms to another of the 3 common games, Journey.

- Adventurous, Peaceful, Solitary
- Comradery, Immersive
- Comradery, Immersive, Peaceful, Solitary, Imaginative
- Immersive
- Immersive; Imaginative
- Peaceful
- Peaceful
- Peaceful
- Peaceful
- Peaceful
- Peaceful, Comradery
- Peaceful, Comradery, Imaginative, Sad
- Quirky, Peaceful
- Solitary, Adventurous, Immersive, Meditative, Mystery

What is most interesting here is not that a preponderance of creators (9 out of 14) described Journey's mood as solely or partially Peaceful. It is that 5 out of 15 creators also found Journey to be solely or partially Immersive, and that there is only one instance of overlap between these groups.

The data from Course B shows similar patterns:

- *Blank* (2 times)
- Adventurous, Immersive
- Adventurous, Immersive, Imaginative
- Adventurous, Peaceful, Comradery
- Adventurous, Peaceful, Imaginative
- Peaceful
- Peaceful
- Peaceful, Adventurous
- Peaceful, Quirky, Solitary
- Solitary

For data creators in Course B who implemented mood, a similar percentage (6 out of 9) found Journey to be solely or partially Peaceful. Fewer creators found Journey to be Immersive (2 out of 9) but there was no overlap between Immersive or Peaceful. In Course B, a larger group (5 out of 9) characterized Journey as Adventurous, and this description occurred whether Journey was also Immersive or Peaceful (as was true with Course A, where Adventurous was associated with both Immersive and Peaceful camps). A single creator in Course B, unique amongst all the suggestions of Journey's mood, found it to be just Solitary. Across Course A and Course B, there were alternate invocations of Solitary (5 uses) and Comradery (also 5 uses)—although one person did use these terms together.

These patterns show that, for this data element, merely combining assigned Mood values and taking the most popular ones is not sufficient to understand how data creators interpret Journey's mood. If you did that with this dataset, you'd describe Journey as Peaceful (15 uses), Adventurous (7 uses) and Immersive (7 uses). It's only when looking at how terms are used in combination that one sees Peaceful and Immersive as indicative of *distinct* interpretations of Journey's mood, and not partially constituative of a *single* interpretation of Journey's mood. We see this in alternate uses of Solitary and Comradery to a lesser degree. These patterns also demonstrate how some judgments appear across creation contexts (a number of creators across A and B describe Journey as only Peaceful) while others vary across them (Adventurous is more salient in Course B, as are blank values; creators that use 4 or more mood values are only associated with Course A).

These kinds of observations do not establish a definitive answer regarding the mood of Journey or any other game. Indeed, such observations suggest that no single interpretation will *ever* account for the diversity of approaches to data creation that appear even in this tiny video game dataset. Instead, these insights make use of data creators' different data implementation decisions—a form of design after design—which are made visible through the context provided by the experimental dataset. If we view data aggregation as part of an interlocking set of design layers, we can imagine approaches to aggregation that enact further possibilities for tracing the subtleties of data creation and learning from them, rather than minimizing or erasing those subtleties.

USING DATA, DESIGNING DATA: AN EXAMPLE

I have presented data as an evolving product of interlocking design decisions regarding infrastructure, collection, and aggregation. In this understanding, data takes on a new character with each use, as it is manipulated to fit its new context. This perspective prompts us to make use of data from a more active position: as always designers of data, never its mere appropriators.

By understanding devices and applications for data collection, access, display, and visualization as infrastructuring in an ongoing, dynamic process of data creation, designers can devise artifacts that negotiate and interrogate their relationships with data, acknowledging the artifact's role in shaping data itself. As one example, Houben and colleagues' Physikit is set of 4 cubes that express ambient data visualizations [14]. Each cube expresses data via a different physical modality: light, air flow, vibration, and rotation. Physikit includes a Web configuration tool for users to create rules that control each cube (so that air flow increases as data rises or that a light turns on when a certain threshold is reached). Physikit presents a clever approach to user-controlled, tangible visualization. But although the Physikit designers describe the project as human-data design, Physikit is conceived as a conduit for data, not enmeshed with it. Physikit was "designed to work with any data," and the data source used in the project field study—environmental data collected with the Smart Citizen sensor kit (SCK)—was presented as an external plug-in. In the field study, the designers' goal was to see how participant households would "use" Physikit cubes to understand SCK data.

In my reading of the findings, participants' activities with the cubes were directed toward understanding not just the data and the cubes, but the relationships between SCK, data, and cubes, as implemented in the household setting. Participants "became increasingly suspicious about the accuracy of the kit" and tried to understand "how the data worked." Such experiences reorient the object of design for projects like Physikit. A design-oriented perspective on data suggests that Physikit is not just the cubes and the configuration tool; it's the implementation of a specific dataset with the cubes and configuration tool, and the mechanisms for negotiating those relations. Physikit is infrastructure for making data, not just visualizing it.

Understanding Physikit and other projects in this way suggests a more active role for design in the data itself, as well as in its access and display. For example, Houben and colleagues describe how parents in one household that used the Physikit devised a visualization rule to turn on the light cube when the noise measured by the SCK reached a certain level, to demonstrate to their children how loud they could be. According to Houben and colleagues, the plan "backfired" when the cube lit up for the mother more than the children, showing that she was the noisy one.

Another interpretation of this situation, however, is that the noise data created by family members aligns differently with perceptions of loudness. The SCK sensors record noise levels in decibels, a measurement that describes a ratio, often of power, between two sounds. Decibels, however, are not the only factor in human perception of loudness. A high-pitched sound may be perceived as louder than a deeper sound. A loud high-pitched sound is also unpleasant, whereas a loud deep sound is often perceived as having greater sonority. But all of this is situational also: the pulsing bass in a neighbor's dance music is terrible when you're trying to sleep but invigorating when you're getting ready to go out for the evening.

In the Physikit, when the light cube turns on, the SCK sensor is always measuring higher decibels, but the loudness might vary—or some hearers might perceive the sound as louder than others. If the children's voices are at a higher frequency than the parents', for instance, the decibel level at which the parents become angry and frustrated is probably less than the decibel level at which the parents perceive themselves to be overly loud. As detected by the SCK sensor and visualized by Physikit, the noise data generated by parents and children may very well be different: that is, when the light goes on and the children are making the noise, they might *be* too loud, but when the light goes on and the parents are making the noise, they might *not* be too loud—according to their ears, anyway—even if the decibel levels are the same.

How might Physikit examine this relationship between data, data creator, and data interpreter? One way might be to position the signal (the light) as an implementation question, not as a statement. The light is a cue to ask: is it too loud now? What are the qualities of sound that make it too loud? For whom are these qualities of sound problematic? In the context of aggregation, the salient question is not "Are we louder than other families?" but "How do different households understand and implement loudness?" The visualization can be a reconfigured as a mechanism to understand, contextualize, and interrogate loudness for particular environments.

Of course, although there is an opportunity for projects such as Physikit to examine particularities of data integration, adopting a design perspective on data does not require such a focus. There is no "using" Physikit without a particular dataset, however, and that dataset will have a different character expressed through Physikit than it will through some other means. We can more explicitly acknowledge and describe this kind of relationship in our design projects. This paper contributes by promoting an understanding of this dual relationship: as data plays a role in the design of devices like Physikit, devices like Physikit play an equal role in the ongoing design of data.

Although we have long known that data is interpretively flexible and inherently situational, this knowing has been surprisingly difficult to reconcile with the doing of design, and with the ways that we explain and document design work. What are we doing *to* data when we are doing things *with* it? This paper provides a framework for describing our own roles in shaping data, even with mundane, innocuous-seeming actions such as selecting a form of quantitative measurement, establishing a protocol for automatic data generation, or implementing a mechanism for combining datasets. When we are defining, collecting, and aggregating data, how are we redesigning it? This paper begins a conversation on how to more clearly articulate this kind of design work.

REFERENCES

1. A. Telier. (Thomas Binder, Giorgio De Michelis, Pelle Ehn, Giulio Jacucci, Per Linde and Ina Wagner.) 2011. *Design Things.* Cambridge, MA: MIT Press.
2. Kirsten Boehner and Carl DiSalvo. 2016. Data, Design and civics: an exploratory study of civic tech. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16)., 2970-2981. http://dx.doi.org/10.1145/2858036.2858326
3. Geoffrey Bowker. 2000. Biodiversity datadiversity. *Social Studies of Science* 30, 5: 643-683.
4. Allyson Carlyle. 2015. The policeman's beard was what? Representation and reality in knowledge organization and description. *Proceedings of the iConference 2015.* http://hdl.handle.net/2142/73642
5. Andy Crabtree and Richard Mortier. 2015. Human data interaction: historical lessons from social studies and CSCW In: P*roceedings of the 14th European Conference on Computer Supported Cooperative Work.* 1-20.
6. Chris Elsden, Bettina Nissen, Andrew Garbett, David Chatting, David Kirk, and John Vines. 2016. Metadating: exploring the romance and future of personal data. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 685-698. http://dx.doi.org/10.1145/2858036.2858173
7. Melanie Feinberg. 2015. Genres without writers: information systems and distributed authorship. In *Genre Theory in Information Studies,* edited by J. Andersen. London: Emerald.
8. Joel E. Fischer, Andy Crabtree, Tom Rodden, James A. Colley, Enrico Costanza, Michael O. Jewell, and Sarvapali D. Ramchurn. 2016. "Just whack it on until it gets hot": working with IoT data in the home. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 5933-5944. http://dx.doi.org/10.1145/2858036.2858518
9. Jonathan Furner. 2010. Philosophy and information studies. *Annual Review of Information Science and Technology,* 44: 161-200.
10. William Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a resource for design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03),* 233-240. http://dx.doi.org/10.1145/642611.642653
11. William Gaver, Andy Boucher, Nadine Jarvis, David Cameron, Mark Hauenstein, Sarah Pennington, John Bowers, James Pike, Robin Beitra, and Liliana Ovalle. 2016. The Datacatcher: batch deployment and documentation of 130 location-aware, mobile devices that put sociopolitically-relevant big data in people's hands: polyphonic interpretation at scale. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 1597-1607. http://dx.doi.org/10.1145/2858036.2858472
12. Charles Goodwin. 1994. Professional vision. *American Anthropologist* 96(3): 606-633.
13. Jane Greenberg. 2009. Metadata and digital information. In Bates, Marcia, and Mary Niles Maack (eds). *Encyclopedia of Library and Information Sciences.* 3rd ed. CRC Press.
14. Steven Houben, Connie Golsteijn, Sarah Gallacher, Rose Johnson, Saskia Bakker, Nicolai Marquardt, Licia Capra, and Yvonne Rogers. 2016. Physikit: data engagement through physical ambient visualizations in the home. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16)., 1608-1619. http://dx.doi.org/10.1145/2858036.2858059
15. International Federation of Library Associations (IFLA). 1998. Functional requirements for bibliographic records (FRBR). http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf
16. Steven J. Jackson and Sarah Barbrow. 2015. Standards and/as innovation: protocols, creativity, and interactive systems development in ecology. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 1769-1778. http://dx.doi.org/10.1145/2702123.2702564
17. Eric Kansa, Sarah Kansa, and Benjamin Arbuckle. 2014. Publishing and pushing: mixing models for communicating research data in archeology. *International Journal of Digital Curation* 9,1: 57-70.
18. Bruno Latour. 1999. *Pandora's hope: essays on the reality of science studies.* Cambridge, MA: Harvard University Press.
19. Jin Ha Lee, Hyerim Cho, Violet Fox, and Andrew Perti. 2013. User-centered approach in creating a metadata schema for video games and interactive media. *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL),* 229-238.
20. Jin Ha Lee, Rachel Clarke, and Andrew Perti. 2015. Empirical evaluation of metadata for video games and interactive media. *Journal for the Association of Information Science and Technology (JASIST).*
21. Donald McMillan, Arvid Engström, Airi Lampinen, and Barry Brown. 2016. Data and the city. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 2933-2944. http://dx.doi.org/10.1145/2858036.2858434
22. Johan Redström. 2008. RE: definitions of use. *Design Studies* 29:4, 410–23.
23. John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal tracking as lived

informatics. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (CHI '14), 1163-1172. http://doi.acm.org/10.1145/2556288.2557039

24. Phoebe Sengers and Bill Gaver. 2006. Staying open to interpretation: engaging multiple meanings in design and evaluation. In P*roceedings of the 6th conference on Designing Interactive systems (DIS '06),* 99-108. http://dx.doi.org/10.1145/1142405.1142422
25. Susan Leigh Star and James Griesemer. 1989. Institutional ecology, "translations" and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-1939. *Social Studies of Science* 19, 3: 387-420.
26. Alex S. Taylor, Siân Lindley, Tim Regan, David Sweeney, Vasillis Vlachokyriakos, Lillie Grainger, and Jessica Lingel. 2015. Data-in-place: thinking through the relations between data and community. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 2863-2872. http://dx.doi.org/10.1145/2702123.2702558
27. Peter Tolmie, Andy Crabtree, Tom Rodden, James Colley, and Ewa Luger. 2016. "This has to be the cats": personal data legibility in networked sensing systems. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (CSCW '16), 491-502. http://dx.doi.org/10.1145/2818048.2819992
28. Cesar Torres, Jasper O'Leary, and Eric Paulos. 2016. LiveObjects: leveraging theatricality for an expressive Internet of Things. In *Proceedings of the 2016 ACM Conference Companion Publication on Designing Interactive Systems* (DIS '16 Companion), 5-8. http://dx.doi.org/10.1145/2908805.2908807
29. Ron Wakkary, Audrey Desjardins, and Sabrina Hauser. 2016. Unselfconscious interaction: a conceptual construct. *Interacting With Computers* 28, 4: 501-520.
30. Patrick Wilson. 1968. *Two kinds of power: an essay on bibliographic control.* Berkeley and Los Angeles: University of California Press.
31. Peter Worthy, Ben Matthews, and Stephen Viller. 2016. Trust me: doubts and concerns living with the Internet of Things. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (DIS '16), 427-434. https://doi.org/10.1145/2901790.2901890
32. Marcia Zeng and Lois Mai Chan. 2009. Semantic interoperability. In Bates, Marcia, and Mary Niles Maack (eds). *Encyclopedia of Library and Information Sciences.* 3rd ed. CRC Press.