Cultural Analytics

ENGL 64.05

Fall 2019

Prof. James E. Dobson

"Data are things"

- Rosenberg starts with Paul Edwards's claim that "data are things."

- Can we ever have abstract data?

- Data are measured, shaped, and stored in material form.

- Digitization, for example, represents objects within the present standards and systems.

Figure 5.4

Punch-card storage cabinets in the main entrance hall of the US National Weather Records Center in Asheville, North Carolina, early 1960s.

Image courtesy of National Oceanic and Atmospheric Administration.

Data is also a Word

- By this, Rosenberg means it has a conceptual history.

- In the eighteenth century, it was possible to talk about the data of belief, of Revelation: "How much more then may we be assured of this in Revelation, where we have all the Data, before us?" (561).

Historicity of Data

- Daniel Rosenberg wants us to understand the historicity of our talk about data.

- Contemporary discourse informs how we imagine, name, and use data.

- Example: "Big Data" belongs to (roughly) our present moment.
  - Very large datasets have existed for a *long time.*
  - Our belief in the value of possible correlations within large data alters how we think about big data.

Abstract or Ideal Standards

1. A standard is any set of agreed-upon rules for the production of (textual or material) objects.
2. A standard spans more than one community of practice (or site of activity). It has temporal reach as well in that it persists over time.
3. Standards are deployed in making things work together over distance and heterogenous metrics.
4. Legal bodies often enforce standards, be these mandated by professional originations, manufacturers' organizations, or the state.
5. There is no natural law that the best standard shall win—QWERTY, Lotus 123, DOS, and VHS are often cited as examples in this context.
6. Standards have significant inertia and can be very difficult and expensive to change.

Bowker and Star, 13-14

Abstract or Ideal Systems

1) There are consistent, unique classificatory principles in operation.
2) The categories are mutually exclusive.
3) The system is complete.

Cultural Analytics

Systems are Contested

- Standards and Classification Systems work in practice—they are practices.
- The result of the struggle with "uncertainty, ambiguity, standardizations, and the practicalities of data quality" (26).
- Systems are objects for cooperation (enable).
- Systems always contain boundary objects (16) between communities.
- Classifications shift historically (19).
- Subject to torque—the stress of misalignment between human subjects and classification (27).

The "Stack"

Data

Standards

System

The Book

Sentence

Paragraph

Chapter

Volume
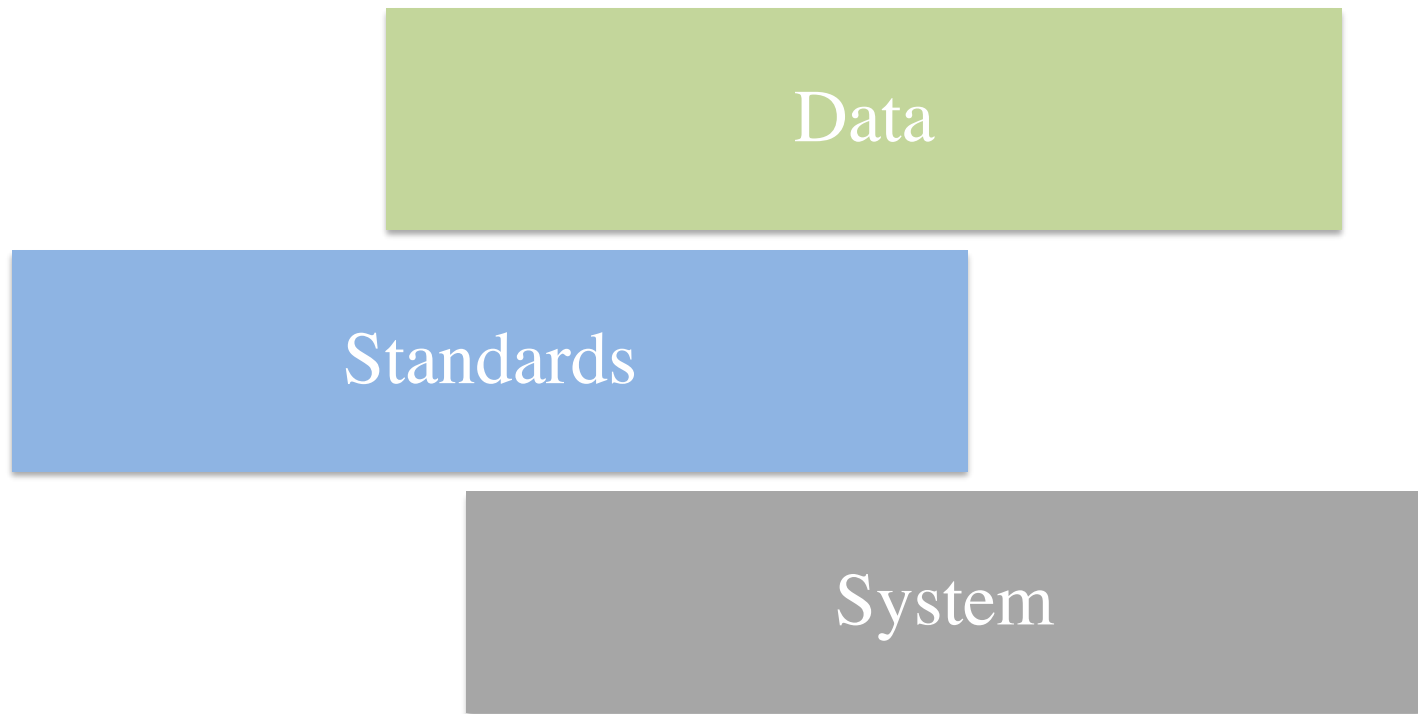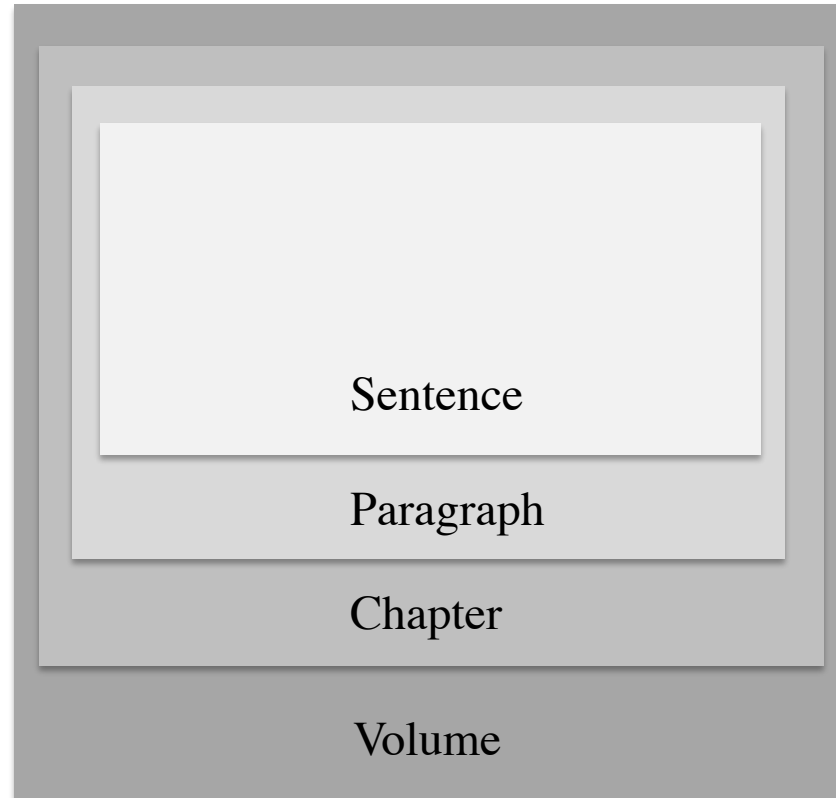
Naming Data

- Downey explains variable assignment as the naming and valuing of an arbitrary identifier.

- Different programming languages and systems (this is data materiality) have different *standards*: different methods and containers.

2.1 Assignment statements

An assignment statement creates a new variable and gives it a value:

```
>>> message = 'And now for something completely different'
>>> n = 17
>>> pi = 3.1415926535897932
```

This example makes three assignments. The first assigns a string to a new variable named message; the second gives the integer 17 to n; the third assigns the (approximate) value of π to pi.

Data vs. Data Type

- Data types are a classification system.
- Downey's example shows us three types: Strings, Integers, Floats.
- Data can be stored in different containers or types—although not all kinds of data can be stored in all containers.
- Python automatically assigns a logical data type.
- Text sources are almost always, at bottom, going to be stored as Strings.
- Word or phrase counts are going to be Integers.
- Higher level statistical data are stored as Floats.
- We can convert from a Float to an Integer but we'll lose *precision*.

Data Types and Affordances

- Additional types enable or *afford* additional operations.

- Converting the type of a variable might enable comparison.

- But sometimes we have complex data.

- How to store publication date of a book, for example?

Type Error as Seen This Morning

Cultural Analytics

Next Class:

- Deconstructing the Raw/Cooked Binary (Lisa Gitelman and Virginia Jackson)
- Gender trouble in the digital archives (Laura Mandell)
- Making Text Data