Defining “Information”

Deepak Kumar

Based (mostly) on Luciano Floridi’s

Information: Three Perspectives

• Information as reality
  (environmental information)

• Information about reality
  (semantic information = data space)

• Information for reality
  (instructions, genetic code, algorithms, etc.)
Information: Perspectives

• Information as reality
  (physical/environmental information)

Information: Perspectives

• Information about reality
  (semantic information)
Information: Perspectives

• Information *for* reality
  (instructions, genetic code, algorithms, etc.)

While TRUE
{
    Code;
}

Information: Three Perspectives

• Information *as* reality
  (physical/environmental information)

• Information *about* reality
  (semantic information)

• Information *for* reality
  (instructions, genetic code, algorithms, etc.)
Information: A Taxonomy

The General Definition of Information

Information = Data + Meaning
The General Definition of Information

Information = Data + Meaning

Question: What is data?
General Definition of a Datum

Dd = \textit{def} x being distinct from y

where the x and y are two un-interpreted variables and the domain is left open to further interpretation.

General Definition of a Datum

datum = \textit{def} x being distinct from y

where the x and y are two un-interpreted variables and the domain is left open to further interpretation.
Understanding Data

What does it mean to...

Damage data
Lose data
Erase Data
Understanding Data...

We have all the data, but no information.

Understanding Data...

Do we now have half the data?
Understanding Data...

Data requires some representation?

Understanding Data...

There is no data?
Understanding Data...

No data is still data!
As long as there is a difference...

General Definition of a Datum

datum =_{def} x being distinct from y

where the x and y are two un-interpreted variables and the domain is left open to further interpretation.
Therefore, this is data even though there is no interpretation and the coding is unknown.
...as long as there is a difference: a lack of uniformity; there is some structure!

Data is...

• “It is the difference that makes a difference.” (McKay 1969)

• A datum is reducible to a lack of uniformity

• The actual format, medium, and language of encoding of data is irrelevant.

• I.e. Information is decoupled from its representation/manifestation.
Data, no Information

<table>
<thead>
<tr>
<th>21.8</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>25</th>
<th>25</th>
<th>25</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.8</td>
<td>1.6</td>
<td>6.84</td>
<td>2.03</td>
<td>6.26</td>
<td>95</td>
<td>114</td>
<td>134</td>
<td>45</td>
</tr>
<tr>
<td>21.8</td>
<td>2.1</td>
<td>8.79</td>
<td>3.6</td>
<td>7.07</td>
<td>160</td>
<td>195</td>
<td>195</td>
<td>60</td>
</tr>
<tr>
<td>21.8</td>
<td>2.4</td>
<td>9.92</td>
<td>4.41</td>
<td>7.72</td>
<td>195</td>
<td>232</td>
<td>217</td>
<td>66</td>
</tr>
<tr>
<td>21.8</td>
<td>3.0</td>
<td>12.1</td>
<td>5.8</td>
<td>9.23</td>
<td>250</td>
<td>279</td>
<td>240</td>
<td>70</td>
</tr>
<tr>
<td>21.8</td>
<td>3.65</td>
<td>14.4</td>
<td>6.82</td>
<td>11.50</td>
<td>292</td>
<td>313</td>
<td>250</td>
<td>75</td>
</tr>
<tr>
<td>21.8</td>
<td>4.25</td>
<td>16.31</td>
<td>7.9</td>
<td>13.61</td>
<td>328</td>
<td>340</td>
<td>258</td>
<td>80</td>
</tr>
<tr>
<td>21.8</td>
<td>4.7</td>
<td>17.74</td>
<td>9.1</td>
<td>14.31</td>
<td>360</td>
<td>364</td>
<td>265</td>
<td>80</td>
</tr>
<tr>
<td>21.8</td>
<td>5.6</td>
<td>20.4</td>
<td>10.65</td>
<td>17.23</td>
<td>383</td>
<td>391</td>
<td>272</td>
<td>80</td>
</tr>
<tr>
<td>21.8</td>
<td>6.45</td>
<td>22.66</td>
<td>11.75</td>
<td>20.52</td>
<td>395</td>
<td>410</td>
<td>270</td>
<td>80</td>
</tr>
<tr>
<td>21.8</td>
<td>7.42</td>
<td>25.36</td>
<td>14.08</td>
<td>22.63</td>
<td>390</td>
<td>425</td>
<td>280</td>
<td>85</td>
</tr>
<tr>
<td>21.8</td>
<td>7.72</td>
<td>26.15</td>
<td>14.35</td>
<td>23.98</td>
<td>390</td>
<td>430</td>
<td>280</td>
<td>85</td>
</tr>
</tbody>
</table>

Common/Useful Representations of Data

- Analog

- Discrete
  (digital)

- Binary
  (y’all know what that is...)
Types of Data

- **Primary**
  (raw data in a database, or a table, ...)

- **Secondary**
  (converse of primary data, stuff that is missing...)

- **Metadata**
  (data about data, location, format, copyright,...)

- **Operational**
  (data about the operation of a data system...)

- **Derivative**
  (data extracted from other data, as in mining...)

The General Definition of Information

**Definition:** $\sigma$ is in instance of information, understood as semantic content, if and only if:

1) $\sigma$ consists of $n$ data, for $n \geq 1$;

2) The data are *well-formed*;

3) The well-formed data are *meaningful*. 
Data are well-formed...

This is an example of data that is not well-formed!
Information: A Taxonomy

- analog
- digital
- binary

- primary
- secondary
- meta
- operational
- derivative

- factual
- instructional
- semantic (content)
- environmental

- data (structured)

- untrue
- intentional (disinformation)
- unintentional (misinformation)

- true (information)
- knowledge
Next,

The Foundations of Information