The Known, Unknown, and Unknowable
How the laws of nature and technology interfere with our needs
Cristian Vava, PhD*

* CEO of Innovatorium Technologies Corporation, Your Special Analytics Department

Innovatorium, ©2010
What is Known Data

Data is known if it is *recorded with acceptable accuracy*

• Known data is always historic
• Recording is from the user’s perspective, what is known to you may be unknown to anyone else
• Accuracy could be very different for each user and application.
  – Measuring temperature within 1% is needless precise for a casual user, just right for the manager of an industrial process, or completely useless for a scientist
  – Same person has different needs at different times.
• Before recording the data the user needs to measure, estimate, or guess its value depending on the desired accuracy and available technology.
Data Measurement

Good accuracy requires to **measure the data** and follow certain procedures to process the results.

Measuring the data implies:
1. Intrinsic knowledge about the environment
2. Clear understanding of the meaning of measured value
Data Estimation

With lax accuracy requirements we could *estimate the data* which involves a mental model for deriving the unknown data from the known.

If the mental model (*observer*) is very well defined mathematically and supported by the natural laws then the estimate could be as good or even better than a direct measurement.

**Example.** A meteorologist may not be able to directly measure the temperature inside of a remote cloud but he considers the temperature known because based on the absorption of radio waves and an appropriate mathematical model he can estimate the temperature with 5% accuracy, better than what he needs.
Data Guessing

If we don’t really care about the accuracy then we could *guess the data*.

Guessing could be an educated guess or a speculation.

Intuition and experience sometimes may allow very good guessing.

To decide how to dress for an outdoor activity all I care about is to guess the outside temperature with accuracy better than 5 °C.
Unknown and Unknowable Data

What is *not recorded* is either *unknown or unknowable*. When the *desired accuracy goes beyond the natural limits* the data becomes *unknowable*.

- If we don’t care much about data’s accuracy the concept of unknowable disappears
- To the other extreme if the accuracy is high then we need to measure the data. Neither guesses not estimations will satisfy us
- Unknown and unknowable could refer to data from the past or the future.
Unknown and Unknowable Data (cont)

To avoid separating the unknown from the unknowable with a fluid and untrustworthy threshold we rely on natural laws. Otherwise due to technological progress what is unknowable today may become a simple unknown tomorrow.

Threshold components

• Natural law limitations like *Heisenberg’s uncertainty principle*
• Other sciences have their accuracy limits imposed by their *own natural laws*
• Impossible to measure data with *arbitrarily high precision*
• *Minimum amount of unavoidable disturbance* caused by the measurement process
Contact Information

Web:  [www.innovatt.com](http://www.innovatt.com)
      [www.heuristicsanalytics.com](http://www.heuristicsanalytics.com)

Email:  bd @ innovatt . com

Phone:  +1-267-342-2815