Why Semantics and Ontology are Important.

Dr. Tom Johnston
MindfulData.com

More generally, a good and appropriate theory provides a set of categories that interpret what you observe, that make sense of it – sometimes without your even being aware of the fact. This is especially true of those parts of a theory which contain terms that are used to describe what one directly observes. The logician W. V. Quine described this part of a theory as lying near the surface of a conceptual sphere. Lying near the heart of that sphere are highly abstract statements such as those of logic and mathematics. In between is the in-between stuff, the stuff where the experts, using their technical vocabularies, extrapolate from their observations, suggest lines for further research, or even make specific predictions.[1]

But contrary to Sir Karl Popper, who believed that single experiments could falsify specific predictions of a theory, Quine has shown that when a theory's predictions prove incorrect, there are many ways for a person committed to the theory to "save the appearances".[2]

The classic example is then, in the early 1600s, the Ptolemaic geocentric theory was being challenged by the Copernican heliocentric theory. Initially, very basic observations seem to confirm the geocentric theory. For example, we seem to see the sun falling below the horizon in the evening, and rising from the horizon in the morning, and to see this as the movement of the sun around the earth. However, observations provided by the newly invented telescope seemed to conflict with the geocentric theory[3]. But that theory's supporters did not roll over and play dead. They introduced "epicycles" to explain the slower or faster than expected motion of various heavenly bodies. With these epicycles, they could "save the appearances" and continue to maintain that the sun revolved around the earth.

My basic point here is that theory *suffuses* observation. It does not just interpret observations made independently of the theory, observations which, because theory-free, are an objective report of what's really "out there". There are no such observations.

---

1 *From a Logical Point of View*, Ch. 1 of *Word and Object*, preface *to Symbolic Logic*.

2 Nelson Goodman. Get the reference.

3 And without a theory of optics, who was to say that what those astronomers saw through those new-fangled instruments was real any more than what we see through a kaleidoscope is real? (This is another illustration of how, often without our being aware of it, theory is embedded in observation, rather than being a separate mechanism to reason about observations made independently of any theory.)

For example, you and I might see squiggly lines under a microscope, where a biologist would see a retro-virus. It's not that the biologist *interprets* squiggly lines as a virus; he *sees* them as a virus, and is confirmed in this way of viewing things by the consensus of his professional community.

So it is with data and data models. A sufficient grounding in semantics and ontology both informs our observations of the data we are to model, and (further towards the center of our conceptual sphere), analyzes and categorizes those observations, and enables us to draw inferences from them that we could not otherwise draw.

One final analogy. When I watch a chess game, I do so with a miniscule but not non-existent understanding of chess openings, end-games and mid-game strategies. But given a move by a chess master, I am never surprised by it. That's because I don't know enough to be surprised. And I certainly don't know enough (past the first two or three moves) to be confident of choosing the best response.

That chess game is a far richer experience for a more informed observer. If one player took sick, and I were to replace him, the game would be lost. If another master replaced him, it might not be. I believe that being able to apply semantics and ontology to data modeling goes a long way towards making one the equivalent of a grand master in chess.

The twin background theories of semantics and ontology are strongly relevant to data modeling. They are also strongly relevant to gathering and clarifying requirements, and in the process clarifying the basic *concepts* which a group of business experts use to describe what they deal with. I have taught subject matter experts to see the data they work with on a daily basis in a new way. I have given them an improved theory about that data, a new ontology. We'll see how this works in Part III, where I apply semantics and ontology to requirements-gathering and clarification.