Finding Meaning in Systems

By Joe Danielewicz,

Joe Danielewicz Consulting, LLC

First published at TDAN.com in October, 2002

Our current definition of, "information," is missing the most important component, namely meaning. Information systems propagate meaninglessness in the form of poorly defined and disintegrated information. This lack of meaning applies to operational type, data warehouse, and business intelligence systems. Part Two of this paper proposes some strategies for managing meaning.

This article will review some of the history of how we arrived at this information vs. meaning crossroads. We will analyze how information systems present meaning, and review some theories of meaning from philosophy.

Information Theory

In February of 2001, Claude Shannon, the father of Information Theory died. He was 84.

In 1948 Shannon was a mathematician employed at Bell Labs working on data encryption when he published, "A Mathematical Theory of Communication.". In his 1948 paper, Shannon states, "the fundamental problem of communication is that of reproducing a message sent from one point, either exactly or approximately, to another point." Shannon's insight was to give precise mathematical definitions to terms like noise, channel capacity, redundancy, entropy, etc. Shannon's theory was so visionary and far reaching, he not only laid the theoretical foundation for the telephony of his day, but his theory applies just as well to optical and wireless communication of today.

The problem with Shannon's Information Theory is that it completely disconnects meaning from information. I suggest it is the lingering after-effects of Information Theory that continue to cause us so much trouble in how we develop information systems.

How many times have you witnessed system developers getting all the requirements right, getting the information through from the database to the user-interface, but not delivering meaning to the business user. The business wants meaning from their information systems, not just accurate communication of messages. The following example illustrates some of the differences.

Information theory tells us that information is anything that makes us feel informed. An example of information that's commonly used is based on something like Monty Hall's, "Let's make a Deal".

Imagine you are presented with three doors and you're told there's a prize behind one of the doors. Your odds of getting a prize are one out of three. Information theory tells us that knowing these odds is information, but this obscures the nature of meaning.

This might make you feel informed, but is this an example of meaningful information? It certainly doesn't tell you anything about what's behind any of the doors. An example of meaning is if each door had a picture (symbols, icons, etc.) that represented what was behind the door. Information theory deliberately obscures the meaning by changing the focus of the communication from meaning to messaging. Information theory ignores the sign relation that is the basis for meaningful statements.

Information Technology's Approach to Meaning: Relational Modeling

Ted Codd and C. J. Date perfected relational algebra for representing facts as relations or tables. Consider the relation, 'CUST'; you could call the attributes of the relation the relational predicate.

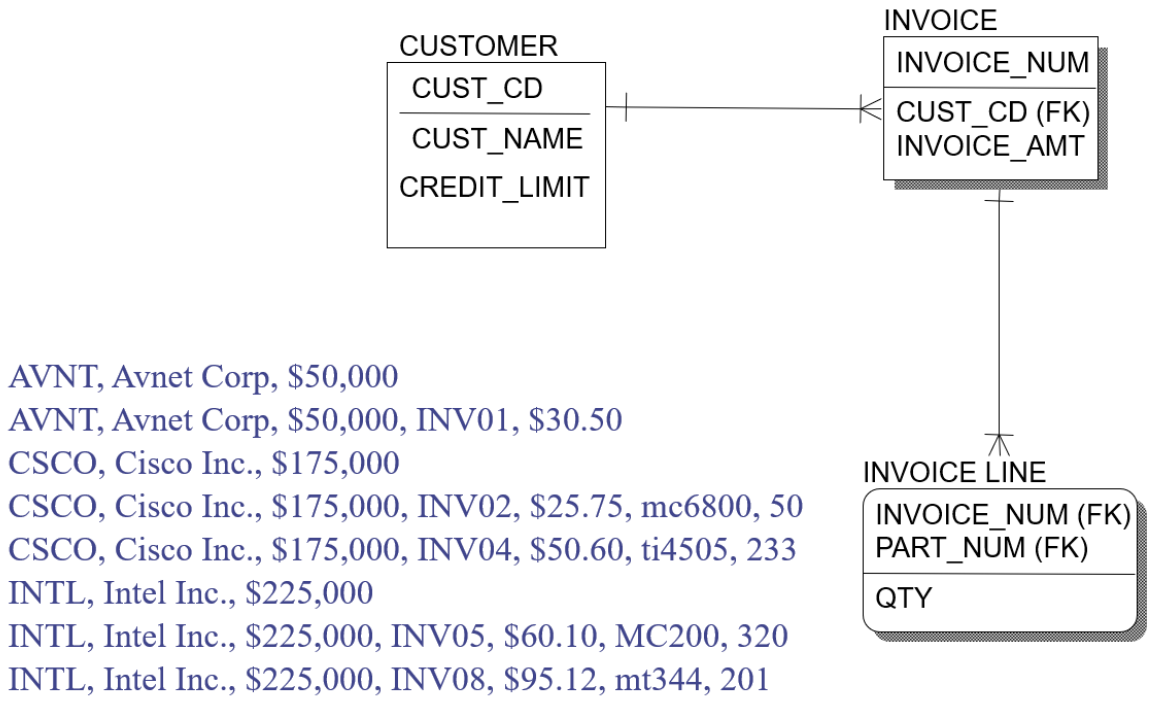

Figure 1 Relational Meaning

By expressing the CUST relation you are asserting some proposition about CUSTOMER INVOICE AND INVOICE LINE.

The base relation can be thought of as the set of true propositions, or facts, about the subject, CUST. The base relation asserts that each of the row instances is a true fact.

Now C.J. Date is quick to point out that a data base management system (DBMS) can't express all of the true facts about the base relation, or topic. But it does express most of the facts that are important to the business.

We just said this base relation doesn't represent ALL true facts about CUST. More than likely there are other relations that have relationships to this table, and all of those propositions are asserted to be true about CUST, also.

The database predicate is the logical 'AND' of the base relation and all the related relations. The relations are equivalent to propositional variables of predicate calculus. And the row instances of these related propositions are asserted to be true statements.

It should be clear that the statements, whether true or false, are different from the facts they are meant to represent. Statements are formed in some sort of artificial or natural language. Statements are true or false. The facts that are represented by statements are a part of the real world; they are neither true nor false, they just are. There's something ungrammatical about a false fact.

This leads us to ask just how does a statement represent a fact? How does language convey meaning?

Information Technology's Approach to Meaning: Metadata Management

Naïve realism and common sense tells us the meanings of words are contained in dictionaries. Information Technology's equivalent of Webster's dictionary is the data dictionary or data repository.
IT's experience with data dictionaries and data repositories is pretty dismal. Business users don't have time to keep looking up meanings in data dictionaries. Context switching to look up a definition back in an on-line dictionary back to their original application is unproductive. Often the dictionary is specific for particular execution environments, such as SAP or Siebel, and looking into these dictionaries is NOT for the business client.

So exactly what does it mean for the data in our systems to mean something, or to have a meaning?

Philosophical literature talks about several theories of meaning.

Referential Meaning

This is the simplest theory of meaning. The meaning of a word is the thing it stands for. This works well for proper names, like the meaning of the word, 'Joe', is the man Joe.
This simple theory doesn't work in all situations. For example, sometimes words refer to things that have no objective existence. What are the objects of reference for mathematical statements? Where do they reside? What does the word, 'unicorn', refer to?

The problem with referential meaning is it's too simple and doesn't cover all instances of meaning. That brings us to the next theory of meaning, the ideational theory.

Ideational Meaning

According to this theory, the meaning of an expression is the thought or idea that is associated with the expression.
This seems to get around the problem with referential meaning; the meaning of the word 'unicorn' is not some real unicorn in the objective world, but my idea of a unicorn. I can have ideas of things that don't exist, and still be able to talk intelligently about them.

The basis for ideational meaning comes from John Locke who believed language is an instrument for the communication of thought. Thought consists of a succession of ideas in consciousness, these ideas being directly accessible only to the thinker. In order to make others cognizant of one's thoughts, one employs publicly observable sounds and marks as representations of these ideas.

One problem with ideational meaning is that when we have a disagreement about the meanings of words, we don't examine people's thoughts to resolve the disagreement. We look in a dictionary.

Stimulus-Response Meaning

This objectivist theory is based on the premise that meaning can be explained by assuming words and expressions are stimuli that elicit certain observable responses. This brings meaning into the domain of observable behavior. The meaning of a chair is, my response to sit down in the chair.

This is loosely based on the B. F. Skinner school of behaviorism. While this theory correctly brings the problem of meaning into realm of social interactions, it falls short of explaining the meaning of meaning.

Meaning in Analytic Philosophy

There were several major insights into theory of meaning in 20th century analytic philosophy and linguistics. This period of philosophy was heavily influenced by Ludwig Wittgenstein who studied under Bertrand Russell and G. E. Moore at Cambridge University.

Wittgenstein is famous for founding two different schools of philosophy. His earlier philosophy was called Logical Atomism which formed the basis for Logical Positivism. His later philosophy became known as Ordinary Language Philosophy.

His Logical Atomism was outlined in his first publication, "Tractatus Logico Philosphicus". The Tractatus is an extremely original and condensed philosophic work. It is arranged as a series of remarks or aphorisms numbered in decimal notation.

Here are some the primary theses of the book:

1. The world is all that is the case.
2. A fact is the existence of states of affairs.
3. A logical picture of facts is a thought
4. A thought is a sentence with a sense.
5. The limits of my language are the limits of my universe.

Wittgenstein later commented on his purpose in the Tractatus, "My whole task consists in explaining the nature of sentences."

Wittgenstein formulated the Picture Theory of propositions. A sentence or proposition is a logical picture of reality. Not just like a picture, but somehow is a logical representation of reality. In many ways this was a foreshadowing of Isomorphic Meaning explained later in this paper.

Wittgenstein's early philosophy assumes there is a universal form of logical propositions, that all correctly formed propositions adhere to, regardless of language. In Information Technology terms, think of this as a corporate-wide DATA DICTIONARY. It's interesting to note Wittgenstein's later philosophy based on ordinary language and language games was a rejection of many of his earlier ideas on language and meaning.

Wittgenstein's ordinary language philosophy was outlined in his book, "Philosophical Investigations", which was published posthumously. Here he rejected the notion of a universal form of language, in favor of "Language Games". These are similar to the rules of any game, for instance the rules of chess, where both parties understand the rules and pieces. Examples of language games are conversations, books, conference presentations, any situation where the players know the rules of communication. It's clear Wittgenstein's language games were meant to apply to more than just the act of speaking; he himself used examples from gesturing to acts of symbolism.

In this context, meaning problems occur when we go outside the rules of the local language game.

The main reason I'm dwelling on this philosophical insight, is because I think it has implications for the success or failure for trying to build huge monolithic data dictionaries, repositories, etc. These types of systems will never capture the richness and expressive flexibility of natural language.

Isomorphic Meaning

Another way to look at meaning is how we process it as a mental state.

In his book, "Godel, Escher, Bach", Douglas Hofstader suggests that meaning comes from the isomorphism between reality and our concepts of reality.

We perceive a pattern in the external world, and the pattern or form of our perception matches our conception or idea or what we perceive. This produces meaning in me.

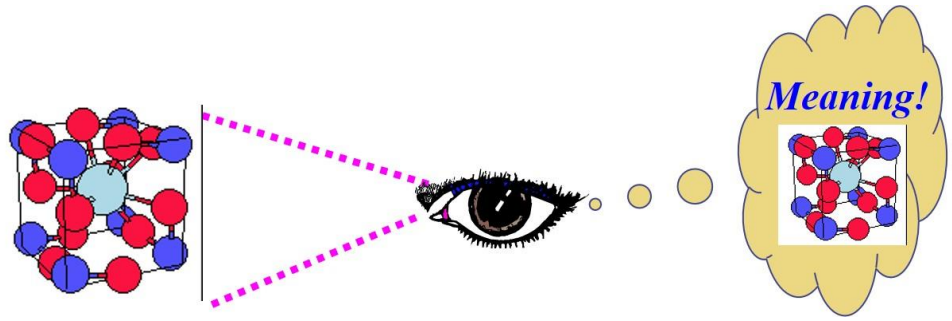

Figure 3 – The logical pattern in my mind matches the pattern I pereive in the external world.

I don't mean the graphical image in my mind matches what I see. I mean the logical pattern in my mind corresponds with reality.

There is something very simple and natural about this explanation because when something means something to me I instinctively feel an alignment between my conception, my words, and reality.

This meaning as isomorphism is similar to the Picture Theory of Propositions from Ludwig Wittgenstein. They both emphasize meaning as an alignment between our concepts and language on the one side, and the reality of the external world on the other side.

Semiotics

Another theory of meaning comes from an American philosopher, Charles Sanders Peirce (1839-1914), founder of pragmatism. Pierce is also considered the father of semiotics, the theory of signs.

A sign, by definition, represents something different than itself. According to semiotics, a sign is a relation consisting of three elements;

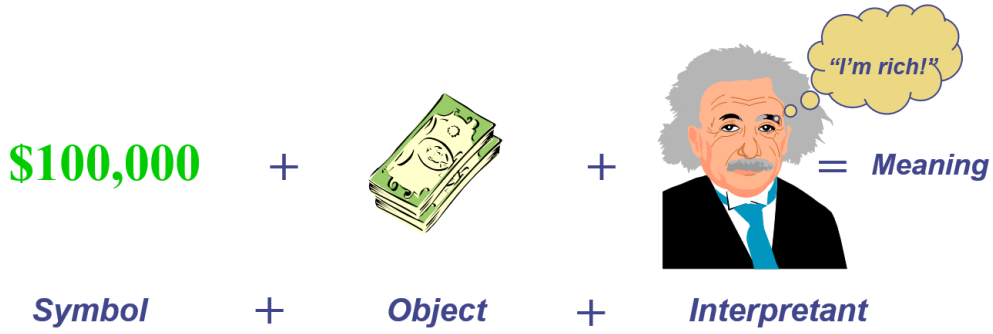

Figure 4 – Semiotic Meaning

Peirce's insight was that meaning depends on all three of these things, sign, referent, and interpretant, standing in relation to each other. Although we can deal with signs, objects and agents independently, when they are in a meaning relation to each other, they depend on each other to create meaning. And the sign relation is recursive in that it creates meaning in an interpretant which can use that meaning as a symbol yet in another sign relation.

Part two: New strategies for managing meaning

Understanding Fallacies in Meaning

Philosophy teaches us about the fallacy of reification. Reification comes from the Latin, 'res', meaning 'thing'. Reification is THING-I-FICATION. Confusion arises when we think of the labels as if they are the real world objects to which they refer. We lose sight of the reality that everything other than the phenomena is an abstraction or symbol.

We speak of PART_NUMBERS as if the data identified by the PART_NUMBER is the real material thing. Then we start ascribing other non-material things to that PART_NUMBER, like CUSTOMER_FORECASTS. Or even worse, we label a CUSTOMER as delinquent, when we really mean, we haven't registered a payment transaction in our accounts receivables.

Or our systems tell us a product is out of stock, when really the quantity-on-hand field in some inventory system has dropped below some threshold. The product maybe in stock but the system doesn't see it.

Situation Theory

Some of the most promising research into meaning and information is being done by the Center for Study of Language and Information (CSLI) at Stanford University. CSLI research spans the fields of language, computation, and cognition and has applications in artificial intelligence (AI). The most interesting area of study relative to meaning is called situation theory.

The goal of situation theory is to develop a unified theory of meaning and information content.

Jon Barwise and Keith Devlin developed a mathematical theory of meaning by explaining meaning as relations between situations and objects like infons and agents. An infon is a discrete item of information. Infons are denoted by

<<P, a1, a2, an, i>>
where **P** is an **n-place** relation, **a**1, **a**2, **a**n, are objects appropriate for the respective argument places of **P**, and **i** is the polarity (0 or 1, also know as true or false) of the statement.
If **i=1** then the infon stands in the relation as stated. In addition to infons and situations, they posit a schema of individuation and an anchor construct which can assign values to parametric infons.
Currently, the research into situation theory is being applied to AI and robotics. Situation theory could be applied to more mundane tasks, such as the user interface (UI) in a business information systems context. This would involve comparing the users profile to the application at hand to determine the relevancy of the infons to the situation.

Develop Data Strategies to Manage Meaning

Exposing all these meaning problems can be really depressing. With all these problems in meaning it's a wonder we ever understand what another person is trying to communicate.

Now what can we do about these problems? How can we really start to manage the meaning of the enterprise?

First, we can understand and be aware of some of these meaning problems. We can understand that the 'data' we all process with our information technology, is not the thing it represents in the external world.

We can use our information technology to develop tools and strategies that promote managing the meaning of the enterprise.

Managing meaning also requires the adoption of different, non-technical, approaches toward system development and deployment. We can probably learn a lot by examining how media enterprises manage meaning. What's common among the dominant content management organizations like, video, music, movie, and other media publishing companies. All of these enterprises have some business strategies in common.

A. Pre-production advertising B. Keep the message simple and short C. Control the distribution method D. The message or information should have a beginning and ending like a story
IT has traditionally controlled the distribution method, but end-user computing and the web have really changed that aspect. IT has never tried to keep their messages short or tried to present the meaning of a system as a story that has a beginning and an end.

This seems to go contrary to establishing a single enterprise-wide standard vocabulary. Don't bother trying to establish a single, monolithic, enterprise wide data dictionary. Business clients

won't let you name their data in a standard manner. They have their own language games that are perfectly valid for their situations. Instead, promote standard language games for various parts of the business. Do this by publishing glossaries and managing the changes to the definitions.

Many companies have data configuration departments that are responsible for managing the data in systems. At some large engineering based companies, this function manages the engineering design documents that describe complex products. Recently, data configuration management has been augmented with departments whose sole responsibility is to train client how to correctly use new IT systems. These groups are sometimes called Change Readiness departments.

Some companies use Change Readiness to define new roles and responsibilities when new IT systems are deployed.

- Promote quality data by appointing data stewards.
- Define steward roles and responsibilities
- Located in business groups supporting maintenance of specific data elements (I.e. yield, cycle time)
- Proactive management of data, via automated metrics, so issues are kept at a minimum.
- Automated business processes are in place, with an escalation process to management, and data stewards take responsibility for the integrity of their data.

Another suggestion is anticipate data and meaning change and help the business client adopt it. This might be as simple as developing user training on new systems, or as complex as managing data cleansing prior to new systems installation.

Here's an example of a Data Readiness Stoplight Chart clients use when putting in a new product definition system.

The GREEN shows the data that has been cleaned, YELLOW shows the data that is being worked, the RED shows the data area that hasn't been started.

Figure 5 Data Clean Up Stop Light Chart

This is managing change to the data culture environment. This is the socialization of language games.

Try publishing language games that describe common business functions; for example, have the Change Readiness group develop narratives around the typical order entry process. For Data Warehouses, design your language games and glossaries around your subject areas, or fact tables. Develop glossaries, scenarios, business rules, SOP's.

Make good use of metrics to measure the things that are important to managing meaning.

And develop your data culture, by focusing the attention of the enterprise on the data meaning not the information delivery.

References

- Buczkowska, Janina. 2001. "Information as the Basis for Representation." Poland: Akademy of Catholic Theology; http://www.bu.edu/wcp/Papers/Lang/LangBucz.htm
- Date. C. J. August 2001. "The Question of Meaning." Business Rule Community, Foundation Matters Commentary. www.BRCommunity.com
- Devlin, Keith. 1991. **Logic & Information.** Cambridge University Press, New York, NY.
- Hofstadter, Douglas R. 1979. ***Godel, Escher, Bach.*** Basic Books, Inc. New York, NY [ISBN 0-465-026854-0].
- Kenny, Anthony. 1973. ***Wittgenstein.*** Harvard University Press. Canbridge, MA [SBN 674-95390-8].
- Lyman, Peter and Hal R. Varian, 2001. "How Much Information." University of California Berkeley; http://www.sims.berkeley.edu/research/projects/how-much-info/summary.html
- Searle, John. 1998. ***Mind, Language and Society.*** Basic Books, Inc. New York, NY. [ISBN 0-465-04519-7].
- Shannon, Claude. E. 1948. "A Mathematical Theory of Communication". Bell System Technical Journal.
- Wittgenstein, Ludwig. 1921. ***Tractatus Logico-Philosophicus.*** Humanities Press Inc. Atlantic Highlands, NJ. [ISBN 0-391-00359-3].