Ontology and Ontologizing – Essential Elements in the Link between Health Data and Value [1]

In computer and information science, an Ontology is a formal representation of a set of concepts and the relationships among them within a domain such as health care. This is how the computer and information system "understand" the data.

So why do they matter?

Why you need an Ontology

If you wish to construct a useful and usable information system that performs complex, broad tasks such as compiling, maintaining, organizing, and presenting a person's health record and driving protocols of care across many providers, for many types of user, then you *must* employ an Ontology. If you don't, then the computer can do little to help. The result will be a disorganized aggregation of data that no one finds useful.

An Ontology can give us a formal representation of meaning in an information system, including the relationships among concepts; create a bridge between the computer and human understandings of the data and relationships; provide a means for understanding the relationships among data from disparate sources; and provide an understood and usable database on which to perform complex tasks.

What is an Ontology made of?

An ontology is a collection ('database') of concepts, such as types of diseases, medications, blood tests, symptoms, and social circumstances, and the relationships between them. For example:

'Influenza'-> 'is a kind of'->'Viral infection'

Influenza and *Viral infection* are concepts and *is a kind of* is a relationship that allows concepts to be formed into hierarchies. Likewise:

'Toe'->'is part of'->'Foot'

A computer can then use a basic set of formal rules ('software') to reason with the concepts and their relationships (sort, select, and generally manipulate them) to do useful things.

The formal rules for manipulating concepts are a kind of algebra for the symbols. The best known are the Description Logics, for which there are now increasingly sophisticated computer tools, driven by the need to build more 'meaning' into the World Wide Web. (*e.g*. find me web pages that show *some kind of Politician in some kind of Embarrassing situation*).

An Ontology links Formal meaning in computers with Human understanding

An ontology imparts formal meaning to the information and builds the bridge between the internal workings of the computer and the external world of human understanding. Within the computer the concepts and relationships are represented by strings of characters (symbols) such as '123XY7DQM9', which it manipulates but are meaningless to people. A language term, such as 'influenza', is attached to a symbol. Although it is meaningless to the computer it allows people to elucidate the intended meaning of the computer's symbol/representation in the real world.

The computer manipulates the symbols according to formal rules and the results are made intelligible to people by converting the outputs into language phrases that appear on the screen. Multiple languages (English, Spanish, etc) can be attached to the same set of formal concepts to create an intrinsically multilingual system.

The quality of the Ontology is measured by how well the internal workings correspond to human understanding. For example if Mrs Smith has influenza then the computer should conclude she has a kind of viral disease and not a kind of broken leg.

An Ontology help us assemble data from diverse sources into a meaningful whole

We create Ontologies to help us build useful and usable information systems. A sufficiently rich Ontology forms an *inter lingua* between the many types of patient data sources and knowledge bases that are found in health care. Systems such as laboratory systems and health plan systems hold and transmit data in their own format. This can be transformed into the concepts represented in an Ontology (blood tests, diseases, medications, etc).

But this 'Ontologizing' is much more than a mere change of format or standardization of the data. Once the many sources and knowledge bases are represented using the same Ontology, the relationships amongst them can be manipulated and a far richer and more complete picture obtained. Useful questions can then be answered such as "show me the information pertinent to this person's diabetes" rather than just "show me the data from St Elsewhere's lab system".

It is analogous to translating related ancient texts into the same language. The relationships between people, events, discoveries, and beliefs come together to form a complete picture of their world. The whole is far greater than the sum of the parts.

Creating a Useful Ontology and Also How to Fail

There have been several very large-scale efforts at constructing large vocabularies that aspire to be formal Ontologies. Practical attempts to make use of them have been remarkably disappointing. Ontologies are not magic. They need to be built and maintained by people. There are several dos and don'ts that are commonly ignored.

Scope: bigger isn't necessarily better

The wider the scope the greater the complexity of an Ontology, and the harder it becomes to build and use the thing. Attempts to create the 'universal ontology for health' for all purposes for all people have all collapsed under their own complexity.

A practical Ontology needs a purpose that is wide enough to be useful but narrow enough to be realizable. This is an empirical matter of judgment and experience, coupled with a clear sense of purpose.

Ontologizing data is not standardizing health care

A common mistake is to confuse building and using an Ontology with standardizing health care. For example ontology developers demand that clinicians agree on a single way of organizing drugs into categories for all purposes for all time before constructing the Ontology and hence the system. This is both unnecessary and impossible. In fact a key role of an Ontology is to cope with the often legitimate real-world diversity in health care. By using an Ontology to 'understand' the data it can be organized in multiple, complementary ways according to need. A sound Ontology can support standardization of practice but doesn't itself demand it. Unnecessary arguments over the 'one true way' destroy projects.

An Ontology is only as valuable as it's applications

An Ontology on its own does nothing for people. Embedding an Ontology in an application releases its value. This is a specialized task that takes significant resources, which developers of 'standard ontologies' always underestimate. All too often the struggle to incorporate the 'general purpose standard' becomes an obstacle to system development rather than an aid.

In Summary

An Ontology:

- is a formal representation of meaning in an information system;
- creates the bridge between the internal world of the computer and the external world of people's understanding;
- provides an *inter lingua* between disparate data sources and knowledge bases;
- allows us to build useful and usable systems for complex tasks in health care.

Remember:

- don't try to divorce the Ontology from its application (the 'universal ontology')
- building and embedding an Ontology in a useful application has pitfalls that require judgment, experience, clarity of purpose, and resources.