From Classical Metaphysics to Medical Information Systems

Barry Smith

Forthcoming in *Humboldt-Kosmos*

An item in the *Wall Street Journal* of 13.11.2001 describes how computers might be used as a means of alerting health officials to problems caused by bioterrorism. Biological attacks, it is noted, will likely be marked by statistical spikes in the purchase of over-the-counter remedies for specific common ailments, and the data pertaining to such purchases can be made available immediately via the computerized inventory systems used by drugstores.

There is a problem in taking advantage of this data however. For while the technology for running databases has reached an impressive state of maturity, the classification systems upon which this technology is based are the products of myriad conflicting *ad hoc* decisions stretching back to the very early days of database design. This means that the data in drugstore computers exists in a bewildering variety of mutually incompatible formats. To resolve these incompatibilities is here a relatively simple matter. In many other such cases, however, the inconsistencies resulting from incompatible classification systems are leading to ever more intractable problems wherever attempts are made to integrate data from different sources – problems reminiscent of the old fable of the Tower of Babel.

Initially the problem of database integration was resolved in case by case fashion as pairs of databases were cross-calibrated by hand. As the numbers and complexity of database systems

increased, however, the idea arose of streamlining these efforts by constructing one single benchmark taxonomy into which all of the various classification systems would need to be translated only once. By serving as a *lingua franca* for database integration this benchmark taxonomy would ensure that all databases calibrated in its terms would be automatically compatible with each other. Interestingly, now, this central classification system was called by information scientists an *ontology*, and it was quickly recognized that work on its construction had more than a few echoes of the metaphysics of old.

The initial premise of the new information systems ontology was that it would be possible to construct a classification system so general that all databases could be reformulated in its terms. The potential advantages of ontology thus conceived are obvious. If all databases can be made compatible, then the prospect arises of using the thousands of person-years of effort that have been invested in creating separate database resources in such a way as to create a single knowledge base of a scale hitherto unimagined, thus fulfilling an ancient dream of a Great Encyclopedia comprehending the entirety of human knowledge.

Unfortunately however, as experience has shown, there are many difficulties standing in the way of the implementation of this dream. The construction of a benchmark ontology proved to be a much more complex task than was originally envisaged. This ontology must be, on the one hand, simple enough that it can be programmed into our computers. On the other hand it must be comprehensive enough to allow the expression of terms derived from all competing systems of classification. In the face of such difficulties the information systems community has responded with a series of partial ontologies, each resting on a different

pragmatically motivated choice about the way an ontology should be built. Ironically, therefore, the very Tower of Babel conditions which the ontological project was initially designed to address have been recreated within ontology itself.

The Institute for Formal Ontology and Medical Information Science (IFOMIS) in the University of Leipzig represents a new approach to solving the problem of ontology. This Institute, which I have founded together with my Leipzig colleagues Barbara Heller and Heinrich Herre, seeks a return to the original idea of a common reference ontology. In contrast to previous efforts, however, which awarded a prime role to the *practical* factor of programmability, IFOMIS will start from the idea that the project of developing a common reference ontology can profit from the *theories* developed by philosophers over two millennia of ontological research. For many of the difficulties faced by information scientists in building an ontology turn out to be identical to the problems with which philosophers have grappled since Aristotle's day. They are problems relating to universals and particulars, properties and relations, events and processes, wholes and parts, boundaries and connectedness, time and causality – problems whose resolution will need to be based on the work of philosophers both past and present as part of any attempt to lay afresh the groundwork for a principled system of ontology.

The IFOMIS ontology will be marked further by the factor of realism. Thus where existing information systems ontology have been based primarily on the strategy of knowledge representation, that is to say on the attempt to represent the *concepts* used by the practitioners within a given domain, IFOMIS, will seek to develop classification systems which correspond to the divisions and interrelations between the entities *on the side of reality itself*.

The work of IFOMIS will not, however, be exclusively philosophical. Its work is directed also at applications in the domain of medical informatics. IFOMIS will employ not only philosophers but also information scientists and medical specialists, and it will draw on existing research efforts in Leipzig under the auspices of the Institute for Medical Informatics, Statistics and Epidemiology (IMISE) and the Competence Network for Malignant Lymphoma (KML-Leipzig).

The domain of medicine has been selected for application purposes not only in light of its intrinsic significance but also because of the ontological challenges which it presents. Medicine calls for an ontology which can allow the simultaneous application to one and the same reality of distinct perspectives (of, for example, doctor and patient, or pharmacologist and geneticist). This is because the entities in such a complex domain can sustain classifications reflecting causally relevant distinctions *at more than one level of granularity*. The medical ontology of IFOMIS must for example have the resources to sustain not merely an anatomical ontology at the level of organs within the structure of the human body, but also cell, protein, gene and molecule ontologies at successively finer resolutions. It must sustain also classifications of *processes* at different resolutions, including the chemical and biological processes taking place inside the body.

Most importantly, the IFOMIS medical ontology will need to comprehend the various types of entities involved in those complex processes we call *clinical trials*. A clinical trial is a controlled experiment in which the effectiveness of a given therapy is measured in systematic fashion in relation to pre-selected groups of patients. As principal testing ground for its methods, IFOMIS will seek to develop a medical domain ontology that is expressive

enough to represent the structures of all the standard types of trials, an ontology comprehending classification systems for therapies, patient populations and outcomes in such a way as to yield standards not only for the representation of trial data but also for the preparation of clinical protocols and of the guidelines which specify procedures for diagnosis and treatment.

IFOMIS thus has a unique opportunity to put philosophical theories to the test empirically. Its ultimate goal is to provide a complete general ontology of the whole of reality. Its proximate goal is to demonstrate how ontological methods can lead to improvements in the domain of clinical trial management and thus to contribute to the wider effort to achieve an improvement in the reliability, efficiency and economic delivery of health care through the rigorous application of the method of clinical trials.