DATA, INFORMATION, KNOWLEDGE: A SEMIOTIC-SYSTEM'S VIEW FOR DATABASE DESIGN

David William Low
Monash University & Victorian Department of Primary Industries
Australia

Abstract

In this paper, the concepts data, information and knowledge are examined and linked with Charles S. Peirce's semiotic categories. The overall aim of the paper is to propose a Peircean semiotic framework that can be applied to database design generally. The more specific ideas developed in the paper are discussed in relation to a database being developed in the area of weed risk assessment at the Victorian Department of Primary Industries (DPI) in Australia. The argument runs as follows: For a database to be used effectively as a learning resource by its target audience(s), a designer needs to distinguish between the concepts data, information and knowledge. These concepts, it is suggested, can be linked with Peirce's 'three grades of clearness', which in turn, are derived from Peirce's triadic categorical framework, that is, his semiotic. Following Peirce, then, it is argued that if the logical role of each categorical concept is muddied, strategic action and organisational learning by the target audience(s) will be made increasingly difficult, if not impossible. Thus, in communicational terms, the author notes first that data falls into the category of Firstness, and as such, it has no meaning at all. In terms of the application examined, weed risk assessment data must be combined with an organisational structure if it is to become information. Information is therefore linked by the author to the category of Secondness – a resisting structure is identified which defines the data's relevance and makes it something that is useable. Along similar lines, information can be put to use where it is deemed necessary, but its strategic value is entirely uncertain. Thus, it is only at the level of knowledge, which is linked by the author to the category of Thirdness, that we can apply information strategically, that is, with a real-world outcome in mind. Thus, it is argued that while each grade of clearness is necessary to database design, it is only at the third grade of clearness, or at the knowledge stage, that a weed risk assessment database can be used effectively to construct and communicate an ongoing community of enquiry around weed risk science.

Introduction

> *I have not succeeded in persuading my contemporaries to believe that Nature also makes inductions and retroductions.*
>
> Charles Sanders Peirce (c.1900)

> *The truth is we have no data. We have a lot of inferences from data, liable to error, and these we have to correct as best we can by putting them together…and finally, if there is any interest in doing so, ascertain what those observation ought to have been.*
>
> Charles Sanders Peirce (1893)

This paper contests a common starting point and guiding assumption in database design. Almost everything written about database design these days takes as its starting point the assumption that data is something that can be isolated and fixed in some way and thereby called 'factual' (cf. Zins, 2007; Tuomi, 2000; Checkland & Holwell, 1998). In starting out this way, however, it is being assumed that in database design, the really important cognitive work has already been done. Further, as the hard intellectual work has already been completed, it is also being assumed that the communication of the stored data (the "facts") can proceed straightforwardly and un-problematically. The database designer's role within this mundane context, then, is merely to make the already determined data accessible to end users via a suitable interface, report, or query form. Thus, in practical terms, the guiding assumption is that it is the end user who must somehow find connections between the facts and render the facts useful, not the designer, nor some aspect of the database design.

The above view is mistaken on several important grounds. First, to communicate anything, that is, to add something *new* to a user's domain of interest, the facts must be considered to be both propositions and elements of Nature. This implicit double reference (the necessary logical entailment of any representation of a *living* fact) creates a need for explanation. Thus, in communicational terms, the ways we use data are always an *amplification* of some earlier reasoning process. Consequently, the real value of data will be in its potential use in the successful negotiation of *future* events (De Tienne, 2006). As such, the facts are theories to be tested.

In this dialogical context, database designers can no longer assume that data can be thought of as 'authorised facts' and simplistically expect the users to accept and incorporate the data into their forward plans 'on trust'. The data is already in question. Thus, the database must somehow incorporate ways to remain responsive to the dissent of the constituency that interacts to constitute it, especially if it is to exist in any ongoing sense. In this view, a database is an enquiry system that must continue to be responsive to the dissent of its constituency, or it does not exist.

Dissent is defined here as a feeling or way of thinking that opposes an accepted viewpoint. Enquiry is defined as a method of communication that aims to find out the truth. The argument is that by viewing a database as a process of enquiry that is reliant on dissent for its being, it becomes possible to simultaneously allow two possibilities: a database can function to keep communication closed to users (through its self-referential autonomy), and connected to users (through its communicational commitment to engaging with dissent).

The above dialogical view of database design suggests that the purpose of enquiry as a whole is something internal or immanent to any database system. In this sense, databases do not 'transfer' or 'transmit' knowledge between social systems, but, rather, they engage with the world via a method that enables a multifaceted recognition of a shared object of enquiry. Thus, a databases' ability to stabilise itself in the face of both internal and external challenges can be founded upon its ability to recognise its own inherent instability and continually adjust itself (to learn and grow) in response to these challenges to its integrity. It is in this context that the following will examine the logic necessary to progress the idea of a *learning database*.

Living Data

But what logical aspects of a database can connect the facts and makes them useful to future users? Is finding ways to facilitate these connecting logical aspects a part of the design problem? To develop answers to these issues we first note that at the most minimal level of analysis data cannot be viewed as isolated facts, for to claim that it is possible to separate the need for explanation away from 'raw data' is already a kind of connection and demands an explanation. Indeed, the improbability of the separation of facts connected is what is demanding an explanation in a database (cf. Peirce, 1960: 7.198). In short, finding the connections that move data forward into future use is an unavoidable component of database design if the design is to serve a communicative function.

The present counter-claim to the conventional view of data, therefore, is that even 'raw data' can never truly be isolated if it is to be communicational. This leads to another possible objection to viewing data as isolated facts. Because facts cannot be separated without reason, and because all data are simultaneously part of reality and represented, a designer who assumes data to be isolatable would create an unfortunate disconnect between nature and the data, that is, if he or she designed within such a context. The claim here, then, is that the predominant view of database design in effect denies the database the status of a *living* entity.

In this paper, it will be argued that the main reason for the current lack of integration between technologies (such as databases) and natural systems is that the database is no longer being thought of as something that "*lives*, and *moves*, and *has its being* in a logic of events" (Peirce, 1976: 4.344). As Fisher (1996) put the issue succinctly, we cannot deal with the consequences of technologies such as genetic engineering the same way we deal with a broken cup. What is needed, Fisher suggests, are methods that enable us to question the very methods we use to grapple with the dislocations such as pollution, and as we shall examine in more detail shortly, invasive plants (i.e., 'weeds').

As already noted, the paper will address the above design disconnection from the point of view of a dialogical, hence *communicational*, point of view (cf. Ransdell, 2002). In this dynamic view, database design involves more than programming a machine to perform internal enactments of purely predetermined deductive routines. Databases that are designed for learning must do more than merely supply a program for users to draw purely deductive conclusions from the data. In the view to be developed here, the database design must, "explicitly address those communication processes that facilitate the creation of shared meaning" (Tuomi, 2000 p. 114).

The paper's argument as a whole can now be stated. It will be claimed that human - environment disconnections can be embodied in design logic, or not, depending on the logic used to guide the design process. Underpinning this general claim is the idea that when the design of human artifacts (methods of interaction and relation) disregard how a living system thinks – and therefore how it responds to our interventions – this disconnection creates what is commonly called an 'environmental problem'. In this sense, human-environmental disconnections arise from a failure to listen and engage with the 'speech of our environment' (Abram, 1996). Further, ecological systems disintegrate when we become unwilling or unable to listen to the way they object to our ways of thinking and interacting with them (Low, 2008). Indeed, if the communicational root cause of an environmental disconnection goes unrecognised, efforts directed toward reintegrating technologies with natural systems will also be misdirected. Misguided technology-based remedial interventions will therefore make the technology-nature disconnections we are concerned with worse (Fisher, 2006). In short, our remedial efforts in relation to environmental dislocations (such as weed invasions) are possibly being handicapped by limited conceptual interpretations of what is really going wrong.

As noted above, designers often have assumed that they already know what nature has said – the conventional view of data in which data can be thought of as separate, fixed or 'raw' units. In such a view of the world there is no need to view data is if it were a flow of "asymmetric dialogically-structured interpretations that necessarily call forth further interpretations" (Ransdell, 2002). The present paper will argue that this in fact just what is called for when we are dealing with natural systems. Database design can either inhibit or enable living interactions. Therefore, following the lead of Ransdell (2002) and Tuomi (2000), the paper argues that what is necessary for the development of a living database involves the inclusion of the communicational logic of social control practices which make discourse more efficient and effective relative to whatever ends it may have, in an amplicative, future sense.

In the process of developing the above polemic, the paper will draw on work being undertaken by the author at the Victorian Department of Primary Industries (DPI). The author has been commissioned to design a 'Weed Risk Assessment Database" to assist in the work of eradicating or controlling weeds on public and private land in Victoria, Australia. The rationale driving this work is significant: invasive plants and animals are currently costing the Victorian economy an estimated two billion dollars in lost agricultural production and environment related expenses annually (DPI, 2008). Putting this issue into the terms developed earlier, a serious disconnection between the reasoning of natural systems and human technologies (i.e., agriculture, forestry and state parks) has arisen and this has created a serious drain on the Victorian economy.

The Categories: Data, Information, Knowledge

There is a widely held belief in the knowledge management literature that data, information and knowledge form a sequential order (Zins, 2007; Working Group for the Prime Minister's Science, Engineering and Innovation Council, 2006; Checkland et al., 1998). Data is assumed to feed into information, and information is assumed to feed into knowledge. In this "linear model", data is seen as the 'basic building block of knowledge' – data is established in basic research, it is then assumed to flow to applied research, then to development, and ultimately to social benefits (Pielke, 2007).

Tuomi (2000) has argued that that the above conventional view of the linear model can be reconsidered in reverse order: data emerges *after* knowledge and information have addressed some practical problem. In Toumi's reversed view, then, data does not become information *after* the addition of meaning, but rather, data is *created* from information by putting information into a predefined structure that completely defines its meaning. Data therefore only exists after certain social, institutional or practical prejudgments have been determined and implemented.

Toumi's reversal of the standard hierarchy is certainly helpful in drawing attention to the problematic identified in this paper, however, it does little to explain the *logical* structure of either the conventional or reversed linear models. Given this, the present paper follows the lead of Barton (2007) and explores whether the terms 'data', 'information' and 'knowledge' can be helpfully linked with Charles S. Peirce's 'three grades of clearness', which in turn, are derived from Peirce's triadic categorical framework, that is, his *semiotic*. Thus, following Peirce and Barton, it is claimed here that data falls into Peirce's logical category of *Firstness*, information falls into the logical category of *Secondness* and knowledge is linked by the author to the logical category of *Thirdness*. Further, just as the information management sciences hold that data, information and knowledge (or reversed cf., Tuomi) form a *system*, so too will it be argued here that each of Peirce's logical categories or grades of clearness is necessary to a system for database design. It will then be argued that it is only at the third grade of clearness, that is, the *logical* level of knowledge, that a database can be used effectively to overcome the person-environment disconnect to communicate an ongoing community of enquiry focussed around a common interest, for example, weed risk.

To place the above claim in context, then, to dissolve environmental dislocations at their deepest level, Fisher (2006) has suggested that three progressively deeper levels, or "generations of awareness" are needed: first, at the *whistle-blowing* level, we establish an awareness of environmental breakdown; next, we formulate and legitimise *political or social insights* into the causes of the environmental breakdown; and, finally, if action at both these former levels fails to resolve the harm, we enter a level of awareness in which we begin to critically examine the very *ways of thinking and questioning* we use to recognise and remedy environmental dislocations. According to Fisher, then, at the deepest level of environmental awareness, we investigate the very world views (values) that generate environmental breakdowns. In other words, we question our methods of reasoning, so these can be reasoned about, and modified. This meta-task, Fisher claims, enables us to dissolve the very conditions that give rise to environmental dislocations.

In this paper, it will be suggested that each of Fisher's levels of concern represent a progressively deeper and more developed interpretation of the ways of thinking and acting that cause environmental dislocations, and each deeper level therefore leads to a distinct style of remedial intervention. In a similar manner, adherents of the conventional hierarchy model of *data*, *information*, and *knowledge* assume that the highest (or deepest) level comes later, after meaning has been added. Thus, and in a similar fashion to Tuomi, a three level, logically entailed model makes it possible to argue that knowledge arises out of the *failure* of solutions attempted at the preceding levels. As noted earlier, data is not fixed. If it truly does represent a living system, it has already moved on.

The important feature of the data-information-knowledge triad in a semiotic view, therefore, is that the *logical* relations between its elements arise in an entailed manner. Thus, in Peirce's logical framework, the most adequate level of understanding (knowledge) presupposes recognition of two lower grades, making the first category of *feeling* (data) the deepest level of contact with the real. Further, for Peirce, all three grades of clearness (e.g., data, information, knowledge) are necessary for attaining an adequate model of human-nature interaction. In other words, for Peirce, each grade of clearness presupposes the attainment of, rather than the failure of, the preceding grade of clearness, making the deepest level of awareness a *direct* (i.e., an experientially derived) recognition of a natural system's being.

The present task has been undertaken within a semiotic context because it is through our knowledge of environmental harm that the need to maintain a socially mediated identification with nature arises (cf. Low, 2008). In order to correct the shared errors in thinking and acting that lead to harm, we must be able to return to the source of our assumptions and revise them in the light of a direct experience with nature. As such, even our most deeply held spiritual assumptions are not held to be axiomatic. Rather, they too are evolving, or as Tuomi puts it, they are "under construction" (2000, p7). If we are to evolve with Nature, then, our thinking must be as free as the thinking of the beings with which we co-exist.

Success in environmental communication (such as with regard to week risk) therefore, should be measured in terms of how well our method of enquiry links us to natural systems, especially to systems that we are unable to observe and interact with unless we discover a method for making such interactions possible. In other words, for Peirce, enquiry is an *erotetic* process: it comprises an interrogatory relation of *both* deep questioning and deep answering. Our interpretations of natural systems help us see what *ought* to be the meaning of our interpretations if their usefulness for resolving environmental dislocations is to be realised. Thus, to design a weed risk database, a system that incorporates responses and ongoing interactions is required. These interactions enable the database to be used to discover the communicational norms that are meaningful within the practical operation of weed risk practice.

The only system that can put remedial actions of the above kind to a test is the system that made the dislocation evident to us in the first place – that is, an ecosystem. The overall purpose in this article, then, is to show how the hierarchical model of knowledge management can be transformed via Peirce's three logical categories to yield a *recursive system for questioning and answering, that is, enquiry.* Each stage (or grade) of the data-information-knowledge model is necessary to the overall process of recognising environmental dislocations, analysing distinctly what it is that makes the dislocation problematic (and therefore amenable to action), and then working to address the issues in a manner in which our thinking and actions are made adequate to the task of collaborating with the being of an ecosystem in a sustainable manner.

It is within the above general context that the paper will next investigate each level of the triadic system of environmental concern for weed risk database design in more detail. The argument, in line with both Fisher and Peirce, is that our interactions with weeds should ideally attain the third grade, or level, of concern, that is, a communicational state of "being in nature", or what we otherwise call knowledge.

1 Data

Data is the realm of Peirce's first grade of clearness, the grade in which simple facts or qualities of *feeling* are prevalent. In this grade, something is recorded by a science, such as botany, and called a fact. Initially, therefore, the data represents an experience linked to a word or idea to which the experience familiarly applies. In the present context, the experiences and ideas are familiar enough and clear enough to be recognized and linked to the term "weed".

Note that in weed science, the factual character of weediness is regarded as a *possible* fact concerning any particular plant or species of plant. For example, Blackberry (*Rubis fruticosus* agg.) may have invaded a farm and reduced the land available for the production of a pasture; hence it is 'a problem.' In making this move, we say that the plant Blackberry possesses the character of weediness. Peirce (CP7.281) stresses that this is a *possible* fact. For example, if Spear Thistle (*Cirsium vulgare*) does not invade the farm, we do not say that it does not possess the characteristic of weediness. Rather, we say it has the potential to be weedy.

As noted earlier, then, facts are always doubly referred to – there is always a difference between the experienced fact and our idea of its character. Thus, what is necessary of the facts with regard to a particular plant species in the data grade of clearness is whether or not the characteristics identified (i.e., weediness as a set of general characteristics) *might* belong to the realm of our future experience of the species. This is weed risk assessment in a nutshell. Weediness is the infinite set of possible factual characteristics of weeds.

Thus, a weed risk database is an infinite set of possible weedy facts (cf. Peirce 1960: 7.283).

In the above sense, and as Douglas and Wildavsky (1982) have also argued, environmental problems such as 'weediness' may at first appear to have been randomly selected. However, a random selection is possible only in the case of a finite collection. As noted above, in the case of weediness, there are an infinite number of plant characters that might potentially be chosen from. We say "potentially" because whether or not a selected weedy characteristic is an imaginary characteristic or a real characteristic is, at the data stage, a largely pointless question. In the first grade of clearness, the data refers to a *potential* existence, and as such, it has no *actual* existence. In other words, the concern has not emerged as a socially or institutionally focussed reality. The data has not yet been connected to anything we already know of that would oppose it and make it real for us. For example, at an organisational level, we may find that there are no procedures to deal with a potential weed threat, because, as yet, the problem does not exist, except in the realm of data – the realm of ideas. The concern, then, in the first grade of clearness, is primarily hypothetical. Weed risk scientists want it understood that they know that something has the *potential* to invade and cause damage.

For example, a current concern in weed science is "sleeper weeds". Sleeper weeds have been defined as, "…invasive plants that have naturalised in a region but not yet increased their population size exponentially" (Groves, 2000) Thus, even though the weed data may have connected a potential to a set of possible future events, for example, climate change (Scott, Batchelor, Ota & Yeoh, 2008), at the data grade of clearness it is impossible to act on the concern – the risk has not been recognised, and therefore legitimised at an organisational, or community level.

In sum, in the data stage of design, the experienced concern is often treated as a personal opinion lacking a shared community reference. If a community of enquiry does not form around the concern, the environmental dislocation that is the object of concern will continue uninterrupted and the whistle-blower will be ignored, or perhaps even actively suppressed (cf. Martin, 2008).

2 Information

It has been noted above that in the first grade of clearness data has the logical status of a possible and is therefore general. Thus, to attain the grade of clearness of information, data needs to gain support via being connected to the natural, technological or social structures that co-create that data and thus are implicated in the domain of concern. Peirce (1997) called this grade of clearness variously "necessary", "deductive", or "mathematical" clearness. For the present purpose, the following description from Peirce is worth quoting at length:

> Among the characters we pay attention to in this mode of argument is whether or not the hypothesis of our premises conforms more or less to the state of things in the outward world…Our inference is valid if and only if there really is such relation between the state of

> things supposed in the premises and the state of the things stated in the conclusion. Whether this really be so or not is a question of reality, and has nothing at all to do with how we may be inclined to think. If a given person is unable to see the connection, the argument is none the less valid, provided that relation of facts really subsists. If the entire human race were unable to see the connection, the argument would be none the less sound, although it would not be humanly clear. (1997, pp, 83-84).

Note that Peirce is not suggesting in his description of this grade of clearness that we *test* the result of any particular formal analysis. That will come later in the knowledge grade. What is important in this grade, however, is that the information be relevant to the state of affairs we are predicting will necessarily be the case given the data. Note also that Peirce admits that the thinking information embodies is not due just to our thinking it, but could potentially exist independently of us, and therefore is independent of the database. Thus, the way Nature thinks is not dependent on our thinking, but rather, Nature's thought is something our ideas must co-exist with to yield useful working models of reality.

In the second grade of clearness, then, the database designer must incorporate design features that enable the subject matter of concern (weeds) to be recognised at a political or policy level by the action representatives of society. To do this, the designer must use methods that link the data to social identities that sustain themselves across time (Tuomi, 1996). The aim is to involve others with the concern by having the structures identified as implicated respond to the claims the potential facts can be used to make (but as noted above, not tested as such). Indeed, the identities created should centre on what are usually called "the facts," and as such, the oppositional method and subsequent responses map out a *definition* of the problem for the designer and the database.

As contested facts of existence, *two* subjects define the sites of resistance so mapped. There is both something with intrinsic value and something reacting against the way the person making a claim in relation to weeds believes the facts to be. Arcioni (2004) puts the matter well in discussing the definition of weeds at a political level, "Conflicts emerge due to the possibility of some plants being detrimental to some interests but valuable to others" (p. 457). Peirce called this contested grade of clearness "the category of *struggle.*", thus dualistic distinctions permeate the second grade of design concern. Indeed, protagonists at this grade tend to frame their concern with either/or arguments (e.g., 'it's either a weed, or not'). Arguments constructed with information therefore usually rely on a rhetorical strategy in which, of two polar choices, one leads to an irreparable environmental loss, the other to a major environmental victory (Cox, 1982).

As Pielke (2008) has also noted, at this second stage then, the scientific and political domains become intertwined. The data are typically seen to be the weaker side of the opposition because they only represent *possible* facts, while the more powerful side appears to occupy the already established side of the conflict – the already known (Low, 1996). The oppositional dualism of information is therefore based on a power relation, and consequently solutions based solely on this grade of clearness may tend to merely stimulate further conflict or remedial "fixes". This, in turn, creates a pattern of displacements in which each subsequent act of opposition simultaneously creates its own locus of resistance, transforming every subsequent victory, or loss, into yet another displacement of the original concern. The concern then becomes centred upon who 'wins' and who 'loses' rather than how we should act on the basis of our ideas in future.

Seen more from the perspective of a database designer, Tuomi (2000) notes that the displacements caused by thinking based in the second grade of clearness lead to increased, or compulsive efforts to provide decision-makers with more data. However, as we have seen earlier, this would merely enlarge the scope of the abstraction covered by the data; the data would merely cover a larger multitude of possible variations. As PMSEIC (2006) have pointed out, this common escape route to generating more data has led to an 'exponential growth in data' (p. 23). However, the expected 'data deluge' is seen by the Prime Minster's advisers as an 'opportunity' rather than as something problematic about the way scientific research is being communicated. The so-called 'opportunity' arises out of the absence of ways to meaningfully engage with the data produced. The 'opportunity' is therefore an ideological realignment of a danger. The real possibility being faced is that we will be swamped with data and fall into a technologically generated policy paralysis.

As Fisher (2006) also notes, actions based in the second grade of clearness therefore only succeed in a very limited sense. As noted earlier, victories 'won' in this stage of clearness necessarily generate a 'loser' (usually Nature). As such, solutions based solely on information usually only serve to displace the object of concern from one physical, temporal, social, or ideological context to another (cf. Fisher, 1996). For example, treating a weed with herbicide may advance agricultural interests if the weed is detrimental to farming (Arcioni, 2004). However, this form of control may merely displace the risk into another domain: for example, it may result in the pollution of watercourses resulting in the eradication of an endangered species. Such displacements are now becoming highly complex and ramified. For example, in the USA, a great deal of effort has gone into the development of a herbicide-resistant Creeping Bentgrass (*Agrostis stolonifera* L.). The aim is to provide golf course managers with a more selective method of weed control. Outside of the boundary of commercial turf maintenance, however, Creeping Bentgrass is considered to be an environmental weed. During a trial, seeds from the herbicide resistant plant blew off the test plot and hybridised with wild grasses nearby. This led to the hybridization of wild grass with pollen from the genetically engineered grass (Pollack, 2006).

While whistle-blowing may succeed in drawing attention to the above kind of issue, and information may help in defining it in political terms, we find that the second grade the concern is structured in communicational terms as a conflict. Indeed, when discussing weed control at a political level, Arcioni (2004) is led to lament, "There is no clear mechanism to resolve the conflicts" (p. 457). The analysis here suggests that this may be because the objective of a conflict is to defeat an opponent rather than mediate a solution that genuinely satisfies both sides reciprocally in a collaborative manner. Put another way, the purpose of a conflict is to make known (i.e., to legitimate) a particular view of Nature by means of persuasive techniques that do not require a direct dialogue or mediated contact with Nature. Politics is not corrupted by this style of sophistical communication; rather, it is often *constituted* by it. In other words, the ideal of deep questioning (i.e., finding out the truth) is abandoned in politics in favour of other ideals, often without any recognition that a slippage has occurred. Thus, a concern for the subject matter (e.g., weed risk) can easily be lost sight of if we rely solely on *information*.

3 Knowledge

Put in terms of Peirce's categories, thus far we have seen that the first grade of clearness (feeling: data) is logically *presupposed* in the second grade of clearness (struggle: information). However, to address the communication problems noted above, we need third grade of clearness – a mediate realm of thoughtful action that presupposes both of the other grades. Put in Peirce's terms, rather than focus on what has been made clear about weed risk, we next examine what *ought* to be the meaning of weed risk if its true usefulness as a concept is to be fulfilled (Peirce's pragmatic maxim). In other words, we evaluate the ideas that represent our concerns by testing them in the real world. Thus, the third grade of clearness incorporates the direct experience of Nature of the first grade, the dualistic political struggles of the second grade, and a grade of concern not covered by feelings or struggle. Peirce (1997) called this third grade of concern "thought," which for the purpose of the present paper we will call *knowledge*.

Note that in the integrated systems framework being sketched here, each logical grade of database design presupposes the previous logical grade. Thus, at the third grade of concern, the designer unites two distinct styles of communicative action within a third mediating grade of critical awareness. Any disparity between the *possible* facts and the structural concerns defined by the dual forces of *existence* identified at the stage of struggle is therefore a motivation for further *enquiry,* the result of which we call *knowledge.* Thus, it is through enquiring into the issue of weed risk as a community of concern that something independent of our social expectations (i.e., the intrinsic value of natural processes) can be brought into a meaningful relation with our technologies, and by extension, our ways of interacting with Nature. These relationships can be experimented with to obtain a truly co-operative relation. In other words, in the third grade of database design, we take the general principles operative in natural systems seriously enough to consider the consequences of taking action on the basis of our reasoning about them.

The distinguishing feature of knowledge, then, is that it is experimental. If we act on the basis of our thinking, we are connecting our feelings to the struggles of existence, and acting "with-in" thought. Thought, then, is never a thing in itself: it represents something else to us, for some purpose.

To illustrate how this might work, let us consider the environmental issue raised by the possibility that legal standing can be granted to environmental objects, such as ecosystems (cf. Stone, 1972). To enquire into this possibility involves not only conceding that an ecosystem can "speak" (Abram, 1996), but also that its speech can be heard by a human. In this sense, the botanical characteristics of plants both are, and represent, the speech of plants. Put another way, weed scientists create data that both is, and represents the voice of Nature. Allowing this as a possible move therefore represents the first grade of concern.

Next, a group of people (say, 'land managers') embody the above data and inform people about the concerns of a specific ecosystem (e.g., "The weeds are displacing endangered plants!"). In this second grade of concern, we draw a kind of diagram of the issue to indicate all that is relevant to our finding out an explanation. For example, we might find that the concern involves a genetically modified plant that has escaped from an experimental trial due to a lack of social safeguards. Note that in doing this purely formal analysis, the forces (natural, social, institutional) that uphold the "unthinkable" are given legal (i.e., social) standing. What was formerly something categorised as "rightless" is in the process transformed into something that can be communicated with (Kevelson, 1996). The second grade of clearness therefore constructs a definition of the issues that surround any concern. This is what has been called here the second grade of *information*.

In the second grade, then, politicising what was considered unthinkable is made doable. Those involved can save themselves from the dilemma this creates (i.e., the dilemma in this case is: worry-free turf maintenance *versus* natural ecosystem integrity) by denying that any connection between human activities and ecosystems exists, or, alternatively (and here we enter the third grade of concern), the wider community can concede the possibility that ecosystems have legal rights, and consider this as a *provisional* conclusion. Several alternative paths of action then lie open to choice and the community is now free to enquire into the possible future consequences of taking the dissenting voice of Nature seriously.

Built into the third grade of concern, then, is a possible *future* relationship between people and Nature. Nature, when seen as "property" in this sense, is a concept involving the negotiation of power, and this is why we can, if we want to, legitimately specify concepts that create a relationship between humans and Nature, such as "privilege," "obligation," or "responsibility" (Kevelson, 1996). In contrast, ideas such as "wealth," "asset," or "resource" are characteristic of second-grade thinking: there is no common obligation implied, only linear extensions of the concept of personal or corporate ownership. Similarly, direct perceptions of possible harm to the regularities of natural systems are characteristic of first-grade thinking, and as I have already explained, these are not taken seriously until given existence by the second-grade of concern.

What does this mean for database design? As Tuomi (2000) notes, at a practical level, designing to meet the third grade of clearness means more than 'putting information on the Web'. A learning database that operates at the third grade of concern needs to address, in its design, a semiotic logic that facilitates the creation of shared meaning. Thus, in adopting triadic logic as a guiding principle, the designer would incorporate into the design three grades of "…services that enable the development of individual or social capabilities for understanding and acting" (Tuomi, 2006, p34, see also Collins, Colvin & Ison, 2009 on natural resource management).

In practice, then, we might say that a database design should value openness. In the present semiotic framework, and following the recent lead of Tuomi (2006), we might then find that it becomes both possible and necessary to identify and implement *three grades of openness*. In the first grade of openness, the design would need to provide non-discriminatory access to *data.* In the second grade of openness, the design would need to provide non-discriminatory access to services that can be used by stakeholders to generate positions or resistance (that is, *information*, the precursor to our capacity to act). Last, in the third grade of openness, the design would need to create systems that can be reflexively contributed to; a necessary addition if any new *knowledge* is to emerge.

Conclusion

To review briefly, it was argued here that Peirce's 'three grades of clearness' can be transposed onto the prevailing data-information-knowledge model commonly used in information management. Thus, in the first grade, the greatest emphasis is placed on *data*. In the second grade, data is given existence in the mode of dualistic struggles and called *information*. Finally, in the third grade, experimental reasoning becomes predominant, which produces *knowledge*. At this third grade of design clearness, therefore, the task was seen to be one of finding ways to foster a co-operative community in which participants investigate or enquire deeply into the consequence of our ideas about Nature.

Having also briefly examined Fisher's three stages of environmental concern, we have also found them to coincide satisfactorily with Peirce's three grades of clarity. The author undertook the present task to show why Peirce and Fisher's process distinctions might have relevance for database design, especially databases that purport to represent biological knowledge. The analogy demonstrated shows how a reality-based approach to database design requires the blending of all three grades or levels of concern into a triadic system of ongoing deep enquiry.

The above task, however, forms only a component of a brief now underway to develop a "Weed Risk Assessment Database" for the Victorian Department of Primary Industries. To commence this task, in this paper the author was interested primarily in addressing a perceived weakness with respect to the current theories used to guide database design. The author also wanted to explore what contribution Peirce's semiotic logic might have in guiding database design.

The next step required is to develop a suitable software platform that will realise the vision of a learning database for weed risk assessment as outlined here. Social networking platforms are currently being assessed for their suitability in performing part of this work. At this point, the author notes that interest in social networking platforms usually centres on non-work related aspects. The social permissions and obligations governing non-work related interaction are primarily personal. However, if we examine social networking technologies from the point of view of a social system concerned with enquiry, a quite different set of social permissions and obligations might be found to be operative (cf. Low, 2000).

Given the above, the issues to be further investigated concern finding suitable computer software tools that will increase the effectiveness of the communicational norms that enable and facilitate successful enquiry. In this respect, Peirce's semiotic theory should prove helpful as Peirce's central philosophical concern was with developing suitable analytical conceptions for just that purpose (Ransdell, 2002). Accordingly, the finding here is that Peirce's semiotic logic offers promise: Society needs to feel nature as if *feeling* nature were something that really matters. Society also needs *structure* in order to embody the concerns we call "environmental problems," that is, to give the thought of Nature existence at a political or legal domain. Most of all, however, both society and Nature need a method for *enquiry* through which our guiding ideas can get tested and corrected, so that they sustain life.

References

Abram, D. 1996. *The spell of the sensuous: Perception and language in a more-than-human world*. New York: Vintage.

Arcioni, E. 2004. What's in a name? The changing definition of weeds in Australia. *Environmental & Planning Law Journal* 21:450-465.

Barton, J. 2007. *Systems thinking: Philosophy, methodology and applications to knowledge management.* Unpublished PhD Dissertation, Department of Management, Monash University, Melbourne.

Checkland, P., & Holwell, S. 1998. *Information, systems and information systems.* Chichester: Wiley.

Collins, K; Colvin, J. and Ison, R. 2009. Building 'learning catchments' for integrated managing: designing learning systems based on experiences in the UK and South Africa. *Water Sciences & Technology* 59(4): 687-693.

Cox, J. R. 1982. The die is cast: Topical and ontological dimensions of the *locus* of the irreparable. *Quarterly Journal of Speech* 68 (3): 227–239.

De Maria, W., and Jan, C. 1997. Eating its own: The whistleblower's organization in vendetta mode. *Australian Journal of Social Issues* 32 (1): 37–59.

De Tienne, A. 2006. Peirce's logic of information, Seminario del Grupo de Estudios Peirceanos, Universidad de Navarra, 28 Sep 2006.

DPI. 2008. Strategic Plan 2008-11. Victorian Department of Primary Industries, Landscape Protection. Melbourne.

Douglas, M., and A. Wildavsky. 1982. *Risk and culture: An essay on the selection of technical and environmental dangers.* Berkeley: University of California Press.

Fisher, F. G. 1996. Transdisciplinary teaching: Overcoming the inevitable dislocation in applying science without regard to its social context. *Praxiology* 4: 157–176.

Fisher, F. G. 2006. *Response ability: Environment, health and everyday transcendence.* Melbourne: Vista.

Groves, R.H. 2000. Sleeper weeds. *Proceedings of the Twelfth Australian Weeds Conference*, pp 632-636. Hobart, Tasmania Sept 1999.

Kevelson, R. 1996. *Peirce, science, signs*. New York: Peter Lang.

Low, D. W. 1996. The greenie genre: Noble saviours or planetary fools? *Australian Journal of Communication* 23 (2): 101–109.

Low, D. W. (2000). The nonviolent principles of scientific communication. *Australian Journal of Communication* 27(3):39-48.

Low, D. W. 2008. Dissent and environmental communication: A semiotic approach. *Semiotica*, 172 (5/5), 42-64.

Martin, B. 2008. Enabling scientific dissent. *New Doctor*, No. 88, pp. 2-5.

Peirce, C. S. 1960. *Collected papers of Charles Sanders Peirce.* Hartshorne & Weiss (Eds.), Cambridge, MA: The Belknap Press of Harvard University Press.

Peirce, C, S. 1976. *The new elements of mathematics.* C. Eisele (Ed.), Volume IV. The Hague: Mouton.

Peirce, C. S. 1997. *Pragmatism as a principle and method of right thinking: The 1903 Harvard lectures on pragmatism.* New York: SUNY.

Pielke, R. A. 2007. *The honest broker: Making sense of science in policy and politics*. New York: Cambridge.

Pollack, A. 2006. *The New York Times.* August 16. Online: http://www.nytimes.com/2006/08/16/science/16grass.html?_r=1

Ransdell, J. 2002. The relevance of Peircean semiotic to computational intelligence augmentation. Online (21 May, 2009): http://www.cspeirce.com/menu/library/aboutcsp/ransdell/ia.htm

Scott J.K., Batchelor K.L., Ota N. & Yeoh P.B. 2008. Modelling climate change on sleeper and alert weeds. Wembley, WA: CSIRO Entomology. Online: http://www.csiro.au/files/files/porh.pdf

Stone, C. D. 1972. *Should trees have standing? Toward legal rights for natural objects.* Los Altos, CA: Kaufman.

Tuomi, I. 1996. The communicative view of organizational memory: Power and ambiguity in knowledge creation systems. *Proceedings of the 29th Annual Hawaii International Conference on Systems Sciences*. January 3-6, Maui, Hawaii. IEEE Computer Society.

Tuomi, I. 2000. Data is more than knowledge: implications of the reversed knowledge hierarchy for knowledge management and organizational memory. *Journal of Management Information Systems*, 16(3): 103-117.

Tuomi, I. 2006. Open educational resources: What use they are and why do they matter. Report prepared for the OECD, October. Online (21 May, 2009): http://www.meaningprocessing.com/personalPages/tuomi/articles/OpenEducationalResources_OECDreport.pdf

Working Group for the Prime Minister's Science, Engineering and Innovation Council (PMSEIC). 2006. From data to wisdom: Pathways to successful data management for Australian science. Working Group for the Prime Minister's Science, Engineering and Innovation Council (PMSEIC). Canberra: Department of Education, Employment and Workplace Relations.

Zins, C. 2007. Conceptual approached for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*. 58(4): 479--493.