Organised complexity, meaning and understanding
An approach to a unified view of information for information science

David Bawden
Department of Information Science, City University London, London, UK

Abstract
Purpose – The paper seeks to outline an approach to a unified framework for understanding the concept of “information” in the physical, biological and human domains, and to see what links and interactions may be found between them. It also aims to re-examine the information science discipline, with a view to locating it in a larger context, so as to reflect on the possibility that information science may not only draw from these other disciplines, but that its insights may contribute to them.

Design/methodology/approach – The paper takes the form of an extensive literature review and analysis, loosely based on the approaches of Stonier, Madden and Bates, and including analysis of both scientific and library/information literature.

Findings – The paper identifies the concept of information as being identified with organised complexity in the physical domain, with meaning in context in the biological domain, and with Kvanvig’s concept of understanding in the human domain. The linking thread is laws of emergent self-organised complexity, applicable in all domains. Argues that a unified perspective for the information sciences, based on Popperian ontology, may be derived, with the possibility of not merely drawing insights from physical and biological science, but also of contributing to them. Based on Hirst’s educational philosophy, derives a definition for the information sciences around two poles: information science and library/information management.

Originality/value – This is the only paper to approach the subject in this way.

Keywords Information science, Information theory, Physics, Biology

Paper type Conceptual paper

Introduction

The seemingly empty space around us is seething with information. Much of it we cannot be aware of because our senses do not respond to it. Much of it we ignore because we have more interesting things to attend to. But we cannot ignore it if we are seeking a general theory of information. We cannot live only by reading and writing books (Brookes, 1980, p. 132).

The purpose of this paper is to outline an approach to a unified framework for understanding the concept of “information” in the physical, biological and human domains, and to see what links and interactions may be found between them. This analysis is used to re-examine the information system discipline, with a view to locating it in a larger context. In turn, this picture is used to reflect on the possibility that information science may not only draw from these other disciplines, but that its insights may contribute to them.

The author is grateful to Jutta Haider, Jack Meadows and Toni Weller for insightful comments.
Because the discussion must necessarily take in some rather technical areas, such as quantum mechanics, thermodynamics, and molecular biology, I have relied to a large extent on “popular” works for literature citation; these will lead the interested reader on to the scientific literature itself. In particular, anyone seeking an accessible account of the background to these issues should consult Davies (1987, 1998), Gell-Mann (1994), Von Baeyer (2004), and Penrose (2004). The lengthy reference list is intended to draw attention to an extensive literature that appears to have been largely overlooked by the information science community.

The nature of the information science discipline

There has been considerable debate in the literature as to the exact nature of the information science discipline, and its place within the wider spectrum of the “information-based” disciplines and professions: see, as recent examples, Hjørland (2000), Warner (2001), Webber (2003), Sturges (2005), Zins (2006), and also, for an earlier perspective, Brittain (1980); see also the article by Weller and Haider (2007) in this issue of Aslib Proceedings.

Webber’s (2003) analysis is somewhat discouraging, for those who believe in the concept of information science per se, at least in the UK, since she finds that, although three departments style themselves “Information Science” – that in itself a small proportion of the 15 (at the time of Webber’s (2003) survey) offering CILIP-accredited courses – only one has a course with that title[1].

A similar input to this issue on this is a consideration of the place of LIS departments in the faculty/school structure of their universities within Europe, established by a survey of such departments (Kajberg and Lørring, 2005; see also Webber, 2003, for a similar earlier analysis). There is a wide variety: most are in humanities faculties, but others are in social science, pure science, informatics/computing, business schools, education, etc. This illustrates the lack of any clear agreement as to where the library/information disciplines “fit” academically. This survey also showed that the only subject to be covered in all library/information programmes, without exception, was information seeking and retrieval, indicating a similar lack of clear agreement as to the necessary content of LIS education. The emergence of the “I-school” phenomenon in the USA, with its formation of integrated departments bringing together the disciplines centred on information, particularly including computing and information systems is an analogous change (Buckland, 2005; Cronin, 2005). For LIS departments in the UK, Webber (2003) draws attention to a similar process of restructuring, with only three of 15 retaining the same departmental name over a five-year period.

Professionally, in both the UK and the USA, there was until recently a distinct professional body representing information science: the Institute of Information Scientists and the American Society for Information Science, respectively. Both of these have changed their identity in recent years, but in very different ways. The Institute has merged with the (UK) Library Association, to form the Chartered Institute of Library and Information Professionals. Its American equivalent has moved in another direction, explicitly including Information Technology. ASIS has become ASIST, and remains entirely separate from the American Library Association. This shows the divergent views of the relations between the various information disciplines (see Webber, 2003, for a more detailed account of this).
To bring some clarity to this picture, it is helpful to use the ideas of the educational philosopher Paul Hirst (Hirst, 1974; Hirst and Peters, 1970; Walsh, 1993). He argues that, since disciplines are closely associated with their knowledge base, we can understand a discipline by understanding its “form of knowledge”. There are, for Hirst, seven main domains or forms of knowledge, defined by the fundamental nature of the knowledge and concepts with which they deal:

1. mathematics;
2. physical sciences;
3. human sciences;
4. literature and the fine arts;
5. morality;
6. religion; and
7. philosophy.

Where a discipline equates to one of these forms, it is what would be regarded as a “pure” academic subject. Hirst also recognises “practical disciplines”, based on one of the forms, but oriented toward solving practical problems. Engineering, for example, would be a practical discipline based on the form of the physical sciences.

Many academic subjects, however, do not align neatly with any form. Rather they are focused on a topic or subject of interest, using any of the forms that are useful in studying and understanding it. Hirst refers to these as “fields of study”; they are typically, though not necessarily, multidisciplinary.

In this sense, it seems clear than the information disciplines are best regarded as a field of study, the focus of which is recorded information. It is generally, though not universally, accepted that their main “form” is that of the human or social sciences, but other forms play a part. Given that the divergence of views about the information professions is just as great as that about the academic disciplines, we might want to say that this field of study underlies not one but several practical disciplines.

The information science discipline would then be understood as:

... a multidisciplinary field of study, involving several forms of knowledge, given coherence by a focus on the central concept of human, recorded information, and underpinning several practical disciplines.

If we accept this, then it is hardly surprising that the field has no unique place in academic structures, nor that there is continual reassessment as to how its practical disciplines relate to each other. In order to attempt a clarification, we shall need to consider the nature of “human recorded information”, the central concept for the field.

The nature of information (for information science)

There has been no shortage of attempts to define or explain the concept of information, as it applies to the information sciences[2] (see, for example, Bates, 2006; Noorlander, 2005; Cornelius, 2002; Losee, 1997; Meadow and Yuan, 1997; Mingers, 1997; Buckland, 1991; Liebenau and Backhouse, 1990; Belkin, 1978). Despite these efforts, there is no generally accepted and agreed understanding, at least in any degree of explicit detail, of the concept.
Furthermore, the concept of “information” appears in different guises, in disciplines far removed from the library/information area, including the physical and biological sciences (Bawden, 2001). The intriguing question which this raises is whether this is simply a consequence of the same word being used to denote different concepts – though clearly with something in common – or whether there is indeed a closer link, at a deep level, between the meaning of the term in these different domains. If the latter is the case, then we might hope that a better understanding of this deeper meaning, and its consequences, might be a source of a deeper “theory of information” for information science. It should also enable richer interactions between information science and the other disciplines for which information is an increasingly central concept.

Three authors in particular have addressed this idea, albeit in very different ways: Tom Stonier, Andrew Madden, and Marcia Bates.

Stonier (1990, 1992, 1997) makes the earliest detailed attempt to unify the concept of information in the physical, biological and human domains. Starting from the concept of information as a fundamental constituent of the physical world:

> Information exists. It does not need to be perceived to exist. It does not need be understood to exist. It requires no intelligence to interpret it. It does not have to have meaning to exist. It exists (Stonier, 1990, p. 22, author’s italics).

he proposes relations between information and the basic physical quantities of energy and entropy. (His most radical proposal, that information may be carried by physical particles, which he terms “infons”[3], has not been supported.) Stonier postulates that a general theory of information may be possible based on the idea that the universe is organised into a hierarchy of information levels. He moves on from this to propose an evolutionary view of intelligence, as a property of “advanced information systems”, from self-replicating crystals, through the intelligence of animal and human individuals and societies, through to machine intelligence and “super-intelligences”, and then to an examination of the origins of meaning in the workings of the brain. He identifies self-organising information processing systems as the “physical roots of intelligence”, based on his conception of information as a basic property of the universe.

Madden (2004) focuses on the biological domain in his evolutionary treatment of information, examining information processing as a fundamental characteristic of most forms of life. He argues that Lamarckian evolution – the idea that characteristics acquired by a biological organism during its lifetime can be passed on to their descendants – while discredited in general biology, may be appropriate for understanding the evolution of human societies, including their information behaviour. Madden proposes, for the first time so far as the author is aware, that insights from the information sciences may be valuable to the supposedly more “basic” sciences, in this case the biological sciences, because of the commonality of the “information” concept.

Bates (2005), seeking like Stonier to reconcile the physical, biological and human forms of information, takes the general definition that: “information is the pattern of organisation of everything”. All information is “natural information”, existing in the physical universe of matter and energy. “Represented information” is either “encoded” (having symbolic, linguistic or signal-based patterns of organisation) or “embodied” (encoded information expressed in physical form), and can only be found in association with living creatures. Beyond this, Bates defines three further forms of information:
The remainder of this paper is aimed at building upon these three approaches, to outline an approach to the expansion of a unified concept of information of relevance to information science. Following Stonier and Bates, this will seek to account for information – perhaps of different kinds – in the physical, biological and human domains, and to allow for the ideas of meaning, understanding, knowledge. Following all three authors, it will assume an evolutionary approach. This, in an information science context, evokes Karl Popper’s “evolutionary epistemology” (Popper, 1979), and Popper’s “three world” model will be used later. Full use of the insights and discoveries of the physical and biological sciences will be made, while, following Madden, we allow for the intriguing possibility that the insights of the library and information sciences may contribute to the development of the physical and biological sciences, in so far as information concepts are involved.

Information in the physical domain

Biological entities, including human beings, are part of the physical world. What is implied here is what Bates terms “natural information”, which is neither “embodied” nor “encoded”; the information which Stonier argues is a fundamental attribute of the physical universe: for a similar perspective, see Landauer (1991).

An understanding of the ways in which “information physics” is developing is important in gaining an realistic understanding of the information concept in this domain. These matters are, however, complex, technical and often controversial, both within the physics community and outside. Only a brief account, with limited references, often to popular sources, will be given here.

Since Stonier presented his thesis – following pioneers of the area such as Brillouin (1956) – the role of information in physics has become much more widely accepted: see Beckenstein (2003), Von Baeyer (2004) and Lloyd (2006) for recent popular overviews, and Siegfried (2000) and Roederer (2005) for more technical discussion; see also Leff and Rex (1990, 2003), for collections of papers on these themes over a long period of time. The new discipline of “information physics”, in which information is regarded as a fundamental feature of the universe – analogous to space, time or energy – has advanced to the extent to which one of its exponents can give as a strapline to an article the idea that “The structure of the multiverse is determined by information flow” (Deutsch, 2002). Smolin (1997, 2002) has suggested that novel physical theories may not even involve the idea of space itself, but rather a web of information, with the most fundamental events being information exchanges between physical processes. The most ambitious working out of this view so far is due to Frieden (2004), who presents a recasting of physical science in information terms, deriving an “information Lagrangian”, a mathematical structure in which fundamental physics is usually now presented, albeit usually in terms of quantities such as energy (Penrose, 2004, chapter 20). Although such views are by no means universally accepted, they are an
indication of the extent of the adoption of an “information perspective” in the physical sciences.

There are three main areas of the physical sciences in which “information” is widely regarded as a particularly important issue:

1. the study of entropy;
2. aspects of quantum mechanics; and
3. the study of self-organising systems.

Other intriguing ideas may be found – for example, the analogy between the well-known social science “principle of least effort”, or Zipf’s Law, familiar in the information sciences (Egghe, 1988, 2005, chapter 1; Buckland and Hindle, 1969; Rousseau, 2005), and the physical “principle of least action” (Moore, 1996) – but, in the absence of detailed study, these have to be regarded as metaphors rather than anything more substantial.

Entropy is a concept emerging from the development of thermodynamics in the nineteenth century, and is in essence a measure of the disorder of a system (Penrose, 2004, chapter 27). Given that order, or organisation, is a quality generally associated with information, a qualitative link between information and entropy is evident. This was put on a quantitative footing by the observations by Brillouin, Wiener and others, that the “information content” or “entropy” in Shannon-Weaver information theory takes the same mathematical form as that of physical entropy, devised by Boltzmann and Planck[4] (see Denbigh, 1981; Roederer, 2005; Von Baeyer, 2004; Leff and Rex, 1990, 2003 for original writings). Information may, therefore, in this sense, be regarded as a kind of “negative entropy”, an indication that it may indeed be a fundamental physical quantity. Indeed, entropy itself may be regarded as a measure of a lack of information about a physical system, another indication of a connection between human knowledge and the physical world.

It is also worth noting that the overall tendency of entropy to increase – the universe tending towards disorder – is regarded as the probable underlying cause of the perception of time passing or flowing, since the laws of physics are generally invariant in time (Gell-Mann, 1994, chapter 15). The recording and deleting of information, in memory and in the form of physical records, activities that reduce entropy locally at the cost of expenditure of energy and increase of entropy in the universe as a whole, may therefore be closely associated with our sense of time (Penrose, 2004, chapter 27).

Quantum mechanics, devised in the first years of the twentieth century, is the most successful physical theory yet developed, in terms of its ability to account accurately for experiments and observations. Interpreting it, however, and understanding what it “means”, is notoriously difficult. Intriguingly for our purposes, many of the interpretations available make some reference to information or knowledge (Penrose, 2004, chapter 29).

One of the major difficulties is that the quantum formalism suggests that physical reality is a strange mingling of quantum states, so that the outcome of any observation is a mix of possibilities. Accounting for why this does not happen in practice is as yet problematic.

The most usual, “Copenhagen”, interpretation of quantum mechanics suggests that the mixture of states drops to only one, accounting for what is invariably observed,
only after the intervention of a conscious observer. This is a problematic situation, leading to difficulties as to what exactly counts as an “observer”. Some attempts to deal with this have led to the idea of an “information gathering and utilising system” (IGUS), which would include, but not be limited to, a human (Gell-Mann, 1994, chapter 11). At all events, this interpretation puts information and knowledge squarely at the centre of this fundamental physical theory.

The physicist John A. Wheeler is generally credited with initiating the trend to regard the physical world as basically made of information, with matter, energy, and even space and time, being secondary “incidental” (Beckenstein, 2003), encapsulated in Wheeler’s dictum: “Tomorrow we will have learned to understand and express all of physics in the language of information”. His ideas, and their influence, are described in detail in the contributions in Barrow et al. (2004), and in particular Davies (2004); see also Von Baeyer (2004) and Davies (1987, 2006). Wheeler has taken this approach farther than most, in insisting that “meaningful information” is necessarily involved, and hence that “meaning”, in the mind of a conscious observer, in effect constructs the physical world: “physics is the child of meaning even as meaning is the child of physics”. This situation Wheeler terms the “participatory universe”, in which conscious observers, gathering information and comprehending its meaning, are a vital part: “in short […] all things physical are information-theoretic in origin and this is a participatory universe”. Wheeler expressed these ideas in two of a series of challenging RBQs (really big questions). The first – “What makes meaning?” – addresses the issue of how meaning emerges from what conventional physics tells us is a meaningless universe. The second – “It from Bit?” – asks whether and how the “its”, the material particles of the universe, arise from their “information-theoretic” origins.

In an alternative way of understanding quantum mechanics, referred to as the “many worlds interpretation”, it is supposed that the many quantum states do not “collapse” to a single reality, but rather continue to exist together. We are only aware of one reality because it is the one that we happen to inhabit; other versions of us inhabit the others. This interpretation does not require the involvement of information or meaning in forming reality, but it does have one remarkable aspect. David Deutsch, an exponent of this view, argues that “knowledge bearing matter”, containing, in Bates’s terms embodied information, is physically special, having a regular structure across the “multiverse”, the totality of these multiple realities (Deutsch, 1997, chapter 8).

Finally, there should be mentioned an approach to quantum mechanics due to the physicist David Bohm (Bohm, 1980; Bohm and Hiley, 1994). This proposed that there is only one reality, and no role is assigned to observers in bringing it into being, but that it is guided into being by a quantum “pilot wave”, determining which of the various possibilities will become reality; Bohn referred this pilot wave as “active information”. Bohm also includes in his theory the idea of an “implicate order” in the universe, existing “folded up” in the physical world, and acting as the basis for emergent order. This idea is closely associated with the idea that the universe is in some way “holographic”, leading again to the idea that information exchange may be a fundamental process (Susskind and Lindesay, 2005; Beckenstein, 2003; Talbot, 1991).

While it is not possible in an article such as this to attempt to deal with the technical issues, the points made above should serve to give the idea that information, organisation, knowledge, and meaning are deeply involved, albeit in different ways, in this fundamental physical theory[5].
Self-organising systems are a topic of relatively recent interest, but are proving to be of importance in a variety of areas in the physical and biological sciences: see Davies (1987, 1998) for popular accounts. The interest in them comes from two perspectives. On the small-scale, it may be observed that simple physical and chemical systems show a propensity to “self-organise”: to move spontaneously towards a mode which is both organised and also highly complex. On the large scale, science must account for the emergence of highly complex organised structures – stars, galaxies, clusters of galaxies, and so on – in a universe which theorists assure us was entirely uniform and homogenous immediately after its creation. It is still not clear what the origins of this complexity are; it is generally assumed to come from gravitational effects, acting on very small inhomogeneities (Davies, 1998, chapter 2). Gravity in the early universe can therefore be seen as “the fountainhead of all cosmic organization […] triggering a cascade of self-organizing processes” (Davies, 1987, p. 135).

The ubiquity of self-organisation has led some scientists to propose that there may be “laws of complexity”, such that the universe has an “in-built” propensity to organise itself in this way; this view is far from generally accepted, but is gaining support:

An increasing number of scientists and writers have come to realise that the ability of the physical world to organise itself constitutes a fundamental, and deeply mysterious, property of the universe. The fact that nature has creative power, and is able to produce a progressively richer variety of complex forms and structures, challenges the very foundation of contemporary science. “The greatest riddle of cosmology”, writes Karl Popper […] “may well be that the universe is, in a sense, creative” (Davies, 1987, p. 5, author’s italics).

… it becomes possible to imagine that a great deal of the order and regularity we find in the physical world might have arisen just as the beauty of the living world would come to be: through a process of self-organisation, by means of which the world has evolved over time to become intricately structured (Smolin, 1997, p. 15).

… [making progress in explaining recent cosmological findings] will require us finding a way to somehow meld general relativity with complexity theory (Kolb, 2007; quoted in Clark, 2007).

The relevance of these issues to information science is that any such complexity laws would be informational in character; that is to say they would act on the information content of the organisation of matter and energy, tending to its increase. This would therefore form the basis of any unified view of information, rooted in its emergence in the physical world.

We should distinguish here between “order” and “organisation” (Davies, 1987, chapter 6). Both are associated with information, but in a rather different way: order may be viewed as a measure of the quantity of information, and organisation as a measure of its quality. Order is, as noted above, associated with the physical quantity of entropy. Organisation is more difficult to quantify: various measures of “logical depth” or “algorithmic complexity” have been suggested, essentially showing to what extent the organisation of the system allows its description to be abstracted and summarised, or “compressed”. (For readable summaries, see Lloyd, 2006; Gribbin, 2004; Davies, 1998, chapter 4; Gell-Mann, 1994, chapter 3; Von Baeyer, 2004, chapter 12[6]. For fuller and more accounts of complexity, see Ellis, 2004; Chaitin, 2005.)
Systems with a high degree of organised complexity are those with the highest amount of “interesting information”. They are intermediate between systems with a high degree of order, but little complexity (for example, an inorganic crystal) and those with a high degree of complexity, but little evident order or organisation (for example, the pattern of leaves on a lawn after a heavy leaf-fall). It is the presence of organised complexity in physical systems that embodies information in the physical world[7]. However, this is as yet information which has no way of being “meaningful”, in any sense. It is worth noting, however, that several of the interpretations of quantum mechanics noted above allowed for some form of involvement of “meaning” in the wider picture.

Information in the biological domain
The “informatisation” of biology has been a remarkable feature of science over the past decades, from the elucidation of the genetic code in 1953 to the sequencing of the human genome exactly 50 years later, and accompanied by a consequent detailed understanding of the ways in which information is passed through generations of living creatures; see Freeland and Hurst (2004) and Dawkins (1995) for concise and lengthy accounts respectively[8].

The concepts of information theory have been extensively applied to biological, and specifically genetic, information from a relatively early stage (Gatlin, 1972). These arguments very much follow on from those relating to the physical world, in terms of increasing levels of organised complexity and information content, the latter generally understood in terms of Shannon’s formalism and its successors (Rosen, 1985; Avery, 2003; Yockey, 2005). The actions of living things in using energy to decrease entropy in their immediate surroundings, by increasing order and organisation developing, reproducing, etc., while increasing the entropy of the universe as a whole in accordance with the second law of thermodynamics, is another informational link between the physical and biological domains[9] (Davies, 1998, chapter 2).

A major change that has come over biology simultaneous with the increasing emphasis on the understanding of genetic information is the tendency to describe life itself as an informational phenomenon. Rather than defining living things, and their differences from non-living, in terms of arrangements of matter and energy, and of life processes – metabolism, reproduction, etc. – it is increasingly usual to refer to information concepts: “the major difference between living and non-living systems is related to the informational organization in the living cell” (Rich, 1980); “if you want to understand life, don’t think about throbbing gels and oozes, think about information technology” (Dawkins, 1986; see also Roederer, 2005).

Life, thought of in these terms, is the example of self-organised complexity \textit{par excellence}. But with life comes a change from the organised complexity in the physical universe: with life we find the emergence of meaning and context. The genetic code, for example, allows a particular triplet of DNS bases to have the meaning that a particular amino acid is to be added to a protein under construction, but only in the context of the cell nucleus[10].

Further, it has become clear that the origin of life itself may best be viewed as an “information event”, as is the subsequent evolution of all life, and the development of intelligence and culture; a number of authors have elucidated this view in various ways (see, for example, Küppers, 1990; Goonatilake, 1991; Plotkin, 1994; Harms, 2004). The crucial aspect is not the arrangement of materials to form the anatomy of a living
creature, nor the beginning of metabolic processes; rather it is the initiation of information storage and communication between generations that marks the origin of life (Davies, 1998, chapters 2 and 3).

Maynard Smith and Szathmáry (1997, 1999) have shown how an “information centred” view of life may be used to treat evolution in terms of biological information. As life has become more complicated, they argue, this is matched by changes in the means by which biological information is coded, stored and transmitted within the living system. They identify a number of “transitions” in the evolution of biological complexity, corresponding to these information changes, including the origin of cells of different kinds, sexual reproduction, multicellular organisms, and animal co-operation and societies, leading ultimately to language and culture.

The record of life on earth is generally agreed to be one of increasing organised complexity. This should not be overstated, and there are biologists who criticise this view for focusing on exceptions (Jones, 2005): species of relatively low complexity still predominate[11] and there are instances of evolution causing a decrease in complexity. Nonetheless, the natural world continues to manifest examples of increasing complexity. This causes discomfort for biologists, for whom it has worrying connotations of purpose or plan. It may, however, be seen as a further example of “complexity laws” applying in the biological, as well as the physical, domain (Davies, 1998, chapter 10; see also Smith and Jenks, 2005).

“Meaning in context” is, of course, not restricted to the biological domain; as will be seen, it underlies all of human recorded information. Here we are dealing with a new development, the origin of consciousness. This is regarded by several of the authors quoted above as a natural extension of the increasing complexity of the biological world; among others who have considered this issue in the same way, though from different perspectives, are Mithen (1996), Penrose (1994), and Donald (1993, 2002).

Information in the human domain

Here we move to the sort of “information” most familiar in the library/information sciences: communicable recorded information, produced by humans in order to convey what Popper terms “objective knowledge”. In addition to the organised complexity and meaning in context of the physical and biological domains, we have conscious participants with an internal mental comprehension of knowledge and an ability to make it explicit.

Even in this familiar domain, there is some controversy about the best way to understand information. Information may be regarded as an entity or “thing” (Buckland, 1991), as a cognitive attribute of the individual mind (Belkin, 1990), or as something created collaboratively (Talja et al., 2006).

There is a particular issue of how information is to be understood to relate to similar entities, most particularly knowledge (see Meadow and Yuan, 1997; Zins, 2006). Two main frameworks to help in understanding this issue are in common use.

One framework, often termed a “scalar” or “pyramidal” model, regards information, knowledge and related concepts as closely related entities which can be transformed into one another, outside the human mind. It is a common sense model, relying on an appeal to the intuitive difficulty of distinguishing between information and knowledge in normal discourse. The usual entities involved are data, information, knowledge and (sometimes) wisdom[12]. Wisdom is usually not addressed directly in information...
science circles (see Rowley, 2006). These are conventionally seen as forming a pyramid – or sometimes a simple linear or scalar progression – with the broad mass of data at the base being distilled to the peak of wisdom. It is pragmatically accepted that the “distillation” process involves what are generally termed “value added” activities: summarising, evaluating, comparing, classifying, etc. Also, moving from data to wisdom is generally seen as setting information within a context, or framework, of existing knowledge, the context giving the meaning. Checkland and Holwell (1998) give a clear account of this viewpoint, adding an additional element, capta – those data to which one pays some interest – between data and information. They see the transformation from capta to information as involving the addition of context, and hence meaning, and from information to knowledge as involving the creation of large structures of related information. Intuitively appealing though it is, this model is far from rigorous; in particular, it is unclear exactly how it is determined when the transition is made between the various states.

An alternative model regards knowledge as something intrinsic to, and only existing within, the human mind and cognition. Knowledge, being subjective, cannot be directly transferred or communicated from one person to another, but must be converted into information first. Information is then regarded as the objective – and therefore communicable and recordable – form of knowledge. Information is thus the bridge between the subjective knowledge in people’s heads. This model is described clearly by Orna and Pettitt (1998). There are some similarities between this view and Popper’s three-world picture, and it falls within a human- or user-centred perspective of information science, while the scalar model is in an information- or system-centred perspective[13].

These conceptual frameworks, though useful, fall short of giving an account of “human information” which may be related to the unified vision being considered here. For this purpose, we need to focus on the entity termed “knowledge” in both the frameworks, as this is the specific and distinctive entity at the “human” level.

This entity has been studied for many centuries by philosophers, under the heading of “epistemology” (see, for example, Goldman, 2003; Moser and Vander Nat, 1987). These studies, however, are of limited value for information science, since they have usually been based on a consideration of what some person knows, generally expressed as an individual’s “justified true belief”. As with the issues of physical information, this is a complex and, in a different sense, technical topic, and justice cannot be done to it here; see Kvanvig (2003) for a detailed treatment. Suffice to say that, from an information science perspective particularly, there are problems with all three elements of this explanation. It does not seem sensible or appropriate to express information, in a library/information context, in terms of what some particular person believes. Justification of information will usually be in terms of what is recorded “in the literature”, rather than in the usual philosophical justification in terms of what some person has experienced. The idea that knowledge should be “true” is particularly problematic, especially if one takes a Popperian view that all human knowledge is imperfect[14]. Floridi (2005), an exponent of a new interest in the “philosophy of information” within the discipline of philosophy itself, recasts the idea of knowledge as “justified, true belief” into the idea that information is “well-formed, meaningful and truthful data”[15]. This seems more suitable for the needs of information science, but does not reflect the rather muddled reality of the human record[16].
Perhaps the most interesting philosophical approach is that of Kvanvig (2003), who argues that we should replace “knowledge” with “understanding” as a focus for interest. Understanding, for Kvanvig, requires “the grasping of explanatory and other coherence-making relationships in a large and comprehensive body of information”. It allows for there to be greater or lesser degrees of understanding, rather than just having knowledge/information or not. Crucially, it allows for understanding to be present even in the presence of missing, inconsistent, incompatible, and downright incorrect, information. It is firmly on the idea of meaning in context, basic to biological information, and therefore underlying human information, which must build on the biological foundation. This seems to be a more appropriate entity than the philosophers' traditional ideas of knowledge, and the author argues elsewhere (Bawden, 2007), that Kvanvig’s “understanding” is a valuable concept for the theoretical foundations of information science[17].

This can be set, as noted above, in a Popperian ontology, with its three “worlds” (Popper, 1979, 1992, chapters 38 and 39; Notturno, 2000, chapter 7):

1. World 1: the physical world, of people, books, computers, buildings, etc.
2. World 2: the internal, subjective mental state of a conscious individual, including their personal knowledge, or understanding in Kvanvig’s terms.
3. World 3: the world of communicable, objective knowledge, or information.

Popper’s ontology, though it was regarded by Brookes (1980) as a suitable philosophical foundation for information science, has been largely overlooked by philosophers, but has recently been recast in detail for the information sciences (Ingwersen and Järvelin, 2005, pp. 48-9) and for the physical/mathematical sciences (Penrose, 2004, chapters 1 and 34)[18]. Ellis (2004) presents an extension of Popper’s ideas to deal with the kind of organised complexity discussed in this article: this extends to four worlds, with human information “relegated” to World 2a. The author would argue that Popper’s ontology, or a variant of it, is still the most appropriate framework for understanding this topic, and useful in practice (Bawden, 2002).

Information in three domains: a unified view?

This paper has argued that information may be seen in the physical domain as patterns of organised complexity of matter and energy, with the added implication that information, and perhaps meaning and consciousness, may underlie and suffuse the physical universe to a remarkable extent. In the biological domain, meaning-in-context emerges from the self-organised complexity of biological organisms. In the human domain, understanding emerges from the complex interactions of World 2, the mental product of the human consciousness, with World 3, the social product of recorded human knowledge. The term “emerges” is used deliberately, for these are emergent properties, that is to say they appear appropriate to their level: physical, biological, conscious and social.

The linking thread, and the unifying concept of information here, is organised complexity. The crucial events which allow the emergence of new properties are:

1. the origin of the universe, which spawned organised complexity itself;
2. the origin of life, which allowed meaning-in-context to emerge; and
the origin of consciousness, which allows self-reflection, and the emergence of understanding, at least partly occasioned when the self reflects on the recorded knowledge created by other selves.

If, therefore, we understood these three origins fully, we would, presumably, understand information itself equally fully, and the ways in which its various forms emerged. Sadly, the beginnings of the universe, of life, and of consciousness, are among the most deep and difficult problems for science (Gleiser, 2004)[19].

It is possible that an informational approach, of the kind being increasingly adopted, may shed light on these origins and transitions, as well as on processes within the three domains. This might take the form of the elucidation of laws of complexity and of self-organisation, operating at a deep level within all three domains. It is not impossible that insights from the information sciences in the human domain could contribute to this. The contribution of information science to other disciplines has been discussed (Cronin and Pearson, 1990; Hahn, 2003), but these have so far not included any impact on the physical and biological sciences. This is the intriguing prospect offered by the unifying vision suggested tentatively here: that a better understanding of the human information domain could contribute to a better understanding of the emergence of complexity and organisation in the biological and physical realms. If this seems far-fetched, and in a sense the “wrong way round”, some of the issues of the role of the consciousness and meaning in quantum mechanics remind us that it may not be an unreasonable prospect to consider.

Conversely, a better understanding of informational issues in the physical and biological domains could shed light on issues of information science. This would certainly not be a reductive approach, as in the discredited approaches to “socio-biology”; rather it would rely on an understanding of similarities and analogies in emergent patterns of complexity and organisation in very different levels and domains.

The information science disciplines revisited
It was suggested earlier in this article that information science could best be understood as:

... a multidisciplinary field of study, involving several forms of knowledge, given coherence by a focus on the central concept of human, recorded information, and underpinning several practical disciplines.

We may now expand this, in light of the issues discussed above, by explaining further the idea of “human, recorded, information” as “a form of organised complexity, providing meaning in context, and promoting understanding”. This can be understood in terms of the Popperian worlds ontology, with the information sciences concerned primarily with the interaction between World 2 and World 3 (Brookes, 1980; Ingwersen and Järvelin, 2005), albeit mediated through World 1 objects. Consideration of this interaction will help elucidate the “several practical disciplines” aspect.

This interaction may be single or multiple. By “single” is meant that what is studied and promoted is the interaction of a single person with World 3, and the World 1 objects and systems in which it is “carried”; this would include such topics as human information behaviour, information seeking, information retrieval, information organisation, information literacy, and so on. By “multiple”, is meant the study and
promotion of the interaction of groups of people, even whole societies, with World 3, and the organisation systems which make this possible: this would include such disciplines as information management, librarianship, records management, archiving, knowledge management, etc.

There will clearly be some overlap between these two in academic terms, for practical purposes such as course planning and some research, but they still form two “poles” for the academic discipline.

The former – individual – pole corresponds to a group of topics identified as a “core” for information science (Bawden, 2007). The latter – multiple – pole corresponds to what Bates (2005) refers to as the “collection sciences”, which bring together objects to aid research, learning, entertainment, etc., the nature of the objects determining the professional discipline. Published recorded information objects are the focus for librarianship, their unpublished equivalents the focus for archiving, records management and document management; information management might claim to cover both, and to be an umbrella concept, non-living and living carriers of embodied information are mainly the province of museums and of zoos and botanical gardens, respectively.

For convenience, and accepting that the terminology will not be acceptable to everyone, the first pole might be termed information science and the second library/information management. The practical disciplines are likely to be associated with the second pole: while we may regret the passing of a historic and honourable term, Information Scientist is no longer a recognised job description outside a few limited contexts. The topics included in the core may be seen as providing the necessary “intellectual underpinning” for the disciplines.

The explanation of the discipline may now be expanded, at the risk of long-windedness to:

A multidisciplinary field of study, involving several forms of knowledge, given coherence by two foci: first on the central concept of human, recorded information – a form of self-organised complexity, providing meaning in context, and promoting understanding – and second on the interaction between Popper’s Worlds 2 and 3; with an intellectual core of information science which underpins several practical disciplines of library/information management.

This way of understanding the discipline seems to bridge the gap between the fundamental issues outlined above, with their potential for highly interdisciplinary research, and the practical needs of those who plan courses and contemplate the future of the information professions.

Epilogue

Information history, as outlined by Weller (2007) in this issue of *Aslib Proceedings*, considers developments over hundreds, or perhaps thousands, of years. The considerations above give a dramatically longer context, into “deep time”, for the origins of biological, and hence human and social, information:

In some as yet ill-understood way, a huge amount of information evidently lies secreted in the smooth gravitational field of a featureless, uniform gas. As the system evolves, the gas comes out of equilibrium, and information flows from the gravitational field to the matter. Part of this information ends up in the genomes of organisms, as biological information [...] Thus all life feeds off the entropy gap that gravitation has created. The ultimate source of biological information and order is gravitation (Davies, 1998, p. 36).
Looking to the future, it is intriguing to speculate that consciousness, intelligence and human information may play a major, and as yet unforeseeable, part in the development of the universe:

As our World 3 products become ever more elaborate and complex [...] so the possibility arises that a new threshold of complexity may be crossed, unleashing a still higher organizational level, with new qualities and laws of its own. There may emerge collective activity of an abstract nature that we can scarcely imagine, and may even be beyond our ability to conceptualize. It might even be that this threshold has been crossed elsewhere in the universe already, and that we do not recognize it for what it is (Davies, 1987, p. 196).

Notes
1. This is the MSc Information Science course at City University London; the first such course in the UK, and, at the time of writing, the only one remaining.
2. The author is not here addressing the wider issue of the “philosophy of information”: for recent overviews of this, see the special issues of Library Trends (Vol. 55 No. 3, 2004) and Journal of Documentation (Vol. 61 No. 1, 2005).
3. Not to be confused with the “infons” introduced by Devlin (1991), as part of an attempt to provide a mathematical theory of information and meaning; Devlin’s infons are abstract entities with his use of situation theory.
4. Information content is essentially the logarithm of the number of different messages possible from the available symbols, while physical entropy is essentially the logarithm of the number of distinguishable ways in which a physical system may be arranged.
5. It may be noted that quantum theory has generated its own formalism of information theory, in which the “bits” of classical information theory are replaced by “qubits”, which may exist in quantum superposition, effectively being 0 and 1 at the same time, and may further be “entangled”, resulting in very unintuitive behaviour. This forms the basis for the developing field of quantum computing: see Nielsen (2002) for a non-technical introduction, and also Deutsch (2004) and Roederer (2005), and for a more technical treatment, see Vedral (2006) and Nielsen and Chang (2000).
6. It may be argued that all of science is a search for such “compressions”, seeking ways of abstracting observational or experimental data through finding patterns, or “laws”, usually through mathematics, the collection of all possible patterns.
7. Some physicists, most notably Smolin, argue that evolutionary self-organisation may provide not only the observable complex patterns of matter and energy, but also physical laws themselves.
8. The “river” in the title of Dawkins’ “River out of Eden”, is the flow of information through generations; a striking metaphor, and one which would have made no sense before the understanding of the informational nature of the basis of life.
9. Those who lament the gap between the culture of science and that of the humanities may be cheered to know that the thermodynamics of living creatures features conspicuously in the early novels of Thomas Pynchon (Simberloff, 1978).
10. It may seem strange to speak of “meaning” when there is no conscious entity in the cell to understand the meaning; however, all that is implied here is that, in a particular context, the information in the DNA code causes a specific and intended response.
11. As Jones (2005, p. 153) puts it, “A billion years ago most organisms were bacteria – and they still are”.
12. This is sometimes referred to as the T.S. Eliot model, from the lines in that poet’s “Choruses from the Rock”: “Where is the wisdom we have lost in knowledge/Where is the knowledge we have lost in information”.

13. Those concerned at dealing with two different models can readily combine them, by considering that the data-capta-information-knowledge-wisdom spectrum refers only to the objective and communicable information transferred between people.

14. This is seen most clearly for scientific knowledge, which is generally subject to incremental, and occasionally revolutionary, advances. Newtonian mechanics, for example, were regarded as inarguably “true” for 200 years, until the insights of general relativity and quantum theory. While the Newtonian is not “untrue”, in the sense that it accords with intuitive experience, and is the basis for most engineering and technology, it is no longer regarded as an adequate account of how the world is. A more specific example is given by our knowledge of the planet Mercury (Sobell, 2005, chapter 3). It was believed for several decades that the length of the day and the year on Mercury were the same, so that the planet had one hemisphere in perpetual sunlight, and the other in perpetual darkness. This view was promulgated in all astronomical literature between the 1880s and the 1960s: “the terms ‘day’ and ‘night’ have no real meaning so far as Mercury is concerned. One hemisphere is perpetually scorched by the solar rays, while no gleams of sunlight ever penetrate to the far side” (Moore, 1955, pp. 41-2). Lurid science fiction stories were written about the exploration of such a strange world, and speculations were made about the possibility of life in the “twilight zone” between the two hemispheres. Later observations showed this view to be mistaken; though the days and nights on Mercury are long, they do exist. It seems a strange perspective to say that there was “no knowledge” about this aspect of Mercury for 80 years, as the current knowledge of the time has been shown to be in error.

15. In this, Floridi notes that he follows a recent trend to equate (human) information with some composite of data and meaning, for example Checkland’s “information equals data plus meaning”.

16. It is particularly at odds with Popper’s concept of objective knowledge; he insisted World 3 must contain incorrect facts and erroneous arguments.

17. Although it is not a part of the argument, it is worth noting that this gives support for the continuing value of “traditional” library/information services in an era of increasingly capable information retrieval systems. An organised collection of appropriate information resources, with services appropriate to the known needs of users, could claim to aid understanding, in a way that a search engine never could.

18. One difficulty with the adoption of this ontology is a lack of clarity – starting with Popper’s own writings – as to exactly what are the constituents of World 3. Popper described it vaguely as “the contents of journals, books and libraries”, but also specified allowed “problems, theories and critical arguments” (including erroneous arguments); he and his followers also allowed numbers and other mathematical objects, music, laws, institutions, etc. Penrose initially restricts it to mathematical truths (“Plato’s world”), but considers allowing in Beauty, Truth and Morality; Ingwersen and Järvelin say simply that it is objective knowledge, and exemplify it by direction signs. Davies (1987) includes social institutions, works of art, literature and religion.

19. They have also been recognised as such for some while: Smolin (1997) notes that, as a young man, he was inspired to study physics by the thought that he might contribute to answering three questions: “What is the universe?”, within the context of the answer to the first question, “What is a living thing?”; within the context of the answers to the first two questions, “What is a human being?”. It now seems likely that both the answers, and their context, may best be given in terms of information.
References


**Corresponding author**

David Bawden can be contacted at: db@soi.city.ac.uk

To purchase reprints of this article please e-mail: reprints@emeraldinsight.com
Or visit our web site for further details: [www.emeraldinsight.com/reprints](http://www.emeraldinsight.com/reprints)