UOC

IN3 Internet Interdisciplinary Institute

www.uoc.edu

Working paper

Data, information and knowledge: have we got it right?

Max Boisot (boisot@attglobal.net)
Researcher (IN3-UOC)
Sol Snider Center for Entrepreneurial Research (The Wharton School, University of Pennsylvania)

Agustí Canals (acanalsp@uoc.edu)
Researcher (IN3-UOC)
Director of Information and Communications Sciences Studies (UOC)

Working Paper Series WP04-002
Submission date: November 2003
Published in: February 2004
Internet Interdisciplinary Institute (IN3): http://www.uoc.edu/in3/eng/index.htm

ABSTRACT

Economists make the unarticulated assumption that information is something that stands apart from and is independent of processors of information and their inherent characteristics. We argue that they need to revisit the distinctions they have drawn between data, information and knowledge. While some associate information with data, others associate it with knowledge. But since few readily associate data with knowledge, this suggests too loose a conceptualisation of the term 'information'. We argue that the difference between data, information and knowledge is in fact crucial. Information theory and the physics of information provide us with useful insights with which to build an economics of information appropriate to the needs of the emerging information economy.

KEYWORDS

Information, Knowledge, Economics of Information, Information Theory, Physics of Information

SUMMARY

To cite this document, you could use the following reference:

BOISOT, Max; CANALS, Agustí (2004). *Data, information and knowledge: have we got it right?* [Online Working Paper. Submission date: 17th November, 2003]. IN3: UOC. (Working Paper Series; DP04-002) [Date of citation: dd/mm/yy]
<http://www.uoc.edu/in3/dt/20388/index.html>

Data, information and knowledge: have we got it right?

MAX BOISOT

IN3, Universitat Oberta de Catalunya, Barcelona, Catalonia, Spain.

Sol Snider Center for Entrepreneurial Research, The Wharton School, University of Pennsylvania.

e-mail: boisot@attglobal.net

AGUSTÍ CANALS[1]

IN3, Universitat Oberta de Catalunya, Barcelona, Catalonia, Spain.

ESADE Business School, Universitat Ramon Llull, Barcelona, Catalonia, Spain.

e-mail: acanalsp@uoc.edu

Version 3, 17th November, 2003

Abstract

Economists make the unarticulated assumption that information is something that stands apart from and is independent of the processor of information and its internal characteristics. We argue that they need to revisit the distinctions they have drawn between data, information, and knowledge. Some associate information with data, and others associate information with knowledge. But since none of them readily conflates data with knowledge, this suggests too loose a conceptualisation of the term ‘information’. We argue that the difference between data, information, and knowledge is in fact crucial. Information theory and the physics of information provide us with useful insights with which to build an economics of information appropriate to the needs of the emerging information economy.

[1] To whom proofs should be addressed. Present address: Av. Tibidabo, 39-43. 08035 Barcelona, Catalonia, Spain. E-mail: acanalsp@uoc.edu

JEL Classification: A12, D20, D80, M21

Key words: Information—Knowledge—Economics of Information—Information Theory—Physics of Information

1. Introduction

Effective cryptography protects information as it flows around the world. Encryption, by developing algorithms that bury information deeply in data, provides "the lock and keys" of the information age (Singh, 1999, p. 293). Thus while the data itself can be made "public" and hence freely available, only those in possession of the "key" are in a position to extract information from it (Singh, 1999). Cryptography, in effect, exploits the deep differences between data and information.

Knowledge and information are not the same thing, either. Imagine, for example, receiving an encrypted message for which you possess the key and from which you extract the following information: "The cat is tired". Unless you possess enough contextual background knowledge to realize that the message refers to something more than an exhausted cat – possibly a Mafia boss, for example - you may not be in a position to react in an adaptive way. To Understand the sentence is not necessarily to understand the message. Only prior knowledge will allow a contextual understanding of the message itself, and the message, in turn will carry information that will modify that knowledge. Clearly, then, information and knowledge must also be distinguished from one another.

In everyday discourse, the distinction between data and information, on the one hand, and between information and knowledge, on the other, remains typically vague. At any given moment, the terms data and information will be used interchangeably; whereas at another, information will be conflated with knowledge. Although few people will argue that knowledge can ever be reduced to data, the two terms are unwittingly brought into a forced marriage by having the term information act as an informal go-between. The growing commercial interest in cryptography, however, suggests innumerable practical circumstances in which the need to distinguish between the three terms is becoming compelling. But if the distinction works in practice, does it work in theory? This is the question that our paper addresses.

Beginning with the second half of the twentieth century, a number of economists – Koopmans, Marschak, Hurwicz, and Arrow – began to concern themselves with the nature of the economic agent as a ”rational information processor”. Since that time, information has become acknowledged as the key generator of wealth in post-industrial societies. We might therefore reasonably assume that, over the past fifty years, mainstream economists, concerned as they are with wealth creation, would have developed a conceptual approach to information that reflected its growing importance to their field.

In this paper, we shall argue that they have some way to go. Both Stiglitz and Lamberton have noted how, even at the end of the twentieth century, the economic profession's conviction that there can be an 'economics of information' still has to reckon with the lack of any consensus as to what specifically it should cover (Stiglitz, 2000; Lamberton, 1998). As Arrow has commented, " It has proved difficult to frame a general theory of information as an economic commodity, because different kinds of information have no common unit that has yet been identified" (Arrow, 1973, p. iii). In fact, Arrow believed that such units were undefinable (Arrow, 1996)[2].Economics, then, is still looking for a way of thinking about information that is adapted both to its own analytical needs as well as to the needs of the emerging information economy.

For this reason, we can support Lamberton's plea that we abandon a unitary and all-purpose concept of information and develop instead a taxonomy based on significant characteristics of information (Lamberton, 1996, pp. xx-xxii). However, descriptions will not, by themselves, build viable taxonomies. Only adequate theorizing will tell us what characteristics will be taxonomically significant. Here we initiate some necessary theorizing that takes as its focus the differences between data, information and knowledge. We shall proceed as follows. First, in the next section (2) we develop a simple conceptual scheme to inform our subsequent discussion. In section 3, we briefly look at how the economic and organizational sciences have dealt with these differences. Both have tended to conflate information and knowledge and to ignore the role of data. In section 4, we examine what information theory adds to the picture. In section 5, we broaden our analysis by introducing

[2] James Boyle has analyzed the incoherence of information economics over a period of fifty years in his *Shamans, Software and Spleens* (1996).

concepts from a new field, the physics of information. Here, the conflation has been of information with data rather than with knowledge – the ”observer” in physics need have no cognitive capacities as such, only a perceptual ability to distinguish between simple physical states. In section 6, with the help of a simple diagram somewhat reminiscent of a production function, we briefly illustrate how the distinction between data, information, and knowledge might be exploited in economic theorizing. In section 7, we explore the implications of our comparative analysis for an economics of information and put forward three propositions. A conclusion follows in section 8.

2. Conceptualizing the Issue

Consider the way in which economists theorize about information in game theory. Game theory deals with a situation in which knowledge is either taken as being asymmetrically distributed or is taken to be common knowledge (Aumann, 1976; Hargreaves Heap and Varoufakis, 1995), the Nash concept specifying both the game’s information requirements and the conditions of its transmission (Kuhn, 1962; Myerson, 1999). These were hardly models of realism. Yet as game theory evolved in the 1980s and 1990s, it imposed ever-less plausible cognitive conditions on economic agents (Binmore, 1990), reflecting its allegiance to neoclassical concepts of information, knowledge, and computability, as well as to the Arrow-Debreu model of Walrasian general equilibrium (Mirowski, 2002).

How, for example, does game theory deal with the situation in which repeated games unfold under dynamic conditions of information diffusion? Here, information is asymmetrically distributed when the first game takes place and is common knowledge by the time the last game occurs. This situation can also be made to work in reverse. Information can start off as common knowledge in a first game and become asymmetrically distributed by the time the last one occurs. Williamson takes this latter outcome as resulting from a "fundamental transformation" wherein an initial large-numbers bargaining process by degrees gets transformed into a small-numbers bargaining process. Here, contract renegotiation involves an ever-decreasing number of players on account of asymmetrically distributed learning opportunities combined with the effects of information impactedness (Williamson, 1985). This second situation might then count as an instance of repeated games in which information gets differentially "impacted" (Williamson’s term) among the different players according to their respective learning abilities as the games unfold to give them anything but “common knowledge”.

How should data, information and knowledge be conceptualized to account for this? Economists struggle. Or not: Hirschleifer and Riley, for example, in their widely read and popular text, *The Analytics of Uncertainty and Information* (1992), hardly deal with definitional issues at all. Taking information to be an input into decision-making, the authors identify the lack of objective information and the lack of information about the future as the key problems they wish to address. A third problem, the limitation of human understanding when dealing with information, the authors choose to ignore on the ground that their intention is to “model economics, not psychology” (Hirshleifer and Riley, 1992, p. 8). Clearly here, the unarticulated assumption – implicitly endorsed by Shannon’s information theory (Mirowski, 2002) - is that information is something that stands apart from and is independent of the processor of information and its internal characteristics. Information itself is loosely defined as either “knowledge”—ie, as a “stock” - or as an “increment to the stock of knowledge”—ie, as “news” or “message”. Like information, knowledge and/or news are assumed to exist independently of a knower or a receiver of news. The tacit assumption that information and knowledge are "things" is widely held. It is, however, a strong assumption, and therefore one that could only follow from an appropriate conceptualization of information, of knowledge, and of the ways in which they relate to each other. Yet nowhere in Hirschleifer and Riley's book is it possible to find a treatment of information and knowledge that is rigorous enough to serve as a basis for such an assumption and for the economic analysis that builds on it.

If Hirschleifer and Riley associate information with knowledge, two other economists, Shapiro and Varian, taking information to be anything that can be digitized, associate it with data (Shapiro and Varian, 1999). Since data is ”thing-like”, it follows that information is also ”thing-like”, a shared property that allows these authors to claim that the new information economy can draw on the same economic laws as those that govern the energy economy. Here again, the way that data and information relate to one another is ignored. Yet, although data might be taken as thing-like and given - that is after all what the roots of the term *datum* (what is given) imply - what is taken to constitute information is always evolving to reflect the changing relationship between agents and data. Thus, whereas the analysis of data lends itself to the application of comparative statics and can be linearized, the analysis of information requires the examination of complex feedback loops and the application of nonlinear dynamics. The view that information is itself a thing rather than a relation points to the survival of essentialist thinking in economics, and of a concern with *being* rather than with *becoming* (Prigogine, 1980).

Since the distinction between data, information, and knowledge is the focus of this paper, we now briefly discuss how it might be approached.

Data can be treated as originating in discernible differences in physical states-of-the-world – that is, states describable in terms of space, time, and energy. Anything that can be discerned at all is discerned on the basis of such differences (Rosen, 1991) and is discerned by agents (Derrida, 1967; Deleuze, 1969). Agents are bombarded by stimuli from the physical world, not all of which are discernable by them and hence not all of which register as data for them. Much neural processing has to take place between the reception of a stimulus and its sensing as data by an agent (Kuhn, 1974). It takes energy for a stimulus to register as data, the amount of energy being a function of the sensitivity of the agent’s sensory apparatus (Crary, 1999). *Information* constitutes those significant regularities residing in the data that agents attempt to extract from it. It is their inability to do so costlessly and reliably that gives encryption its power and that makes the distinction between data and information meaningful. For if data and information were the same thing, the effective encryption of messages – ie, the concealing information in data in such a way that third parties cannot extract it – would be impossible (Singh, 1999).

What constitutes a *significant* regularity, however, can only be established with respect to the individual dispositions of the receiving agent. Information, in effect, sets up a *relation* between in-coming data and a given agent. Only when what constitutes a significant regularity is established by *convention,* can information appear to be objective – and even then, only within the community regulated by the convention. Finally, knowledge is a set of expectations held by agents and modified by the arrival of information (Arrow, 1984). These expectations embody the prior situated interactions between agents and the world - in short, the agent's prior learning. Such learning need not be limited – as required by the theory of rational expectations (Muth, 1961) – to models specifically relevant to the expectations to which they give rise.

To summarize, we might say that *information is an extraction from data that, by modifying the relevant probability distributions, has a capacity to perform useful work on an agent's knowledge base.* The essential relationships between data, information and knowledge are depicted in Figure 1. The diagram indicates that agents operate two kinds of filters in converting incoming stimuli into information. Perceptual filters first orient the senses to certain types of stimuli that operate within a given physical range. Only stimuli passing through this initial filter get registered as data[3]. Conceptual filters then extract *information-bearing* data from what has been so registered. Both types of filters get ”tuned” by the agents' cognitive and affective expectations (Clark, 1997; Damasio, 1999), shaped as these are by prior knowledge, to act selectively on both stimuli and data.

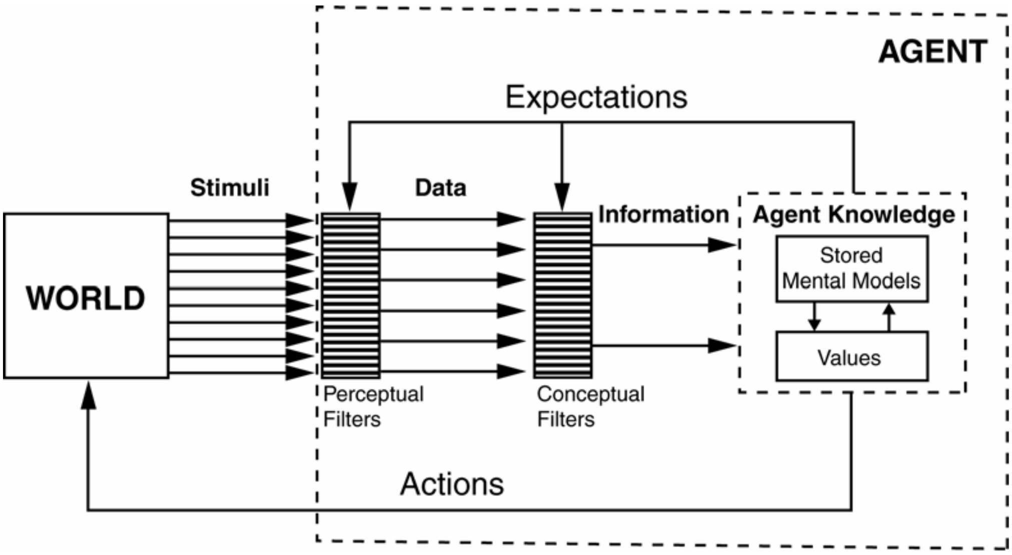

Fig. 1. The Agent-in-the-World

The schema depicted in figure 1 allows us to view data, information and knowledge as distinct kinds of economic goods, each possessing a specific type of utility. The utility of data resides in the fact that it can carry information about the physical world; that of information, in the fact that it can modify an expectation or a state of knowledge; finally,

[3] Roland Omnes, the philosopher of quantum mechanics, understands data thus: "In order to understand what a measurement is, it would be helpful first to make a distinction between two notions that are frequently confused: an experiment's (concrete) data and its (meaningful) result. The data are for us a macroscopic classical fact: thus when we see the numeral 1 on the Geiger counter's screen, this is the *datum.* The *result* is something different, for it is a strictly quantum property, almost invariably pertaining only to the microscopic world, meaning that a radioactive nucleus disintegrated, for example, or providing a component of a particle's spin. The datum is a classical property concerning only the instrument; it is the expression of a fact. The result concerns a property of the quantum world. The datum is an essential intermediary for reaching a result." (Omnes, 1999, author's italics).

that of knowledge in the fact that it allows an agent to act in adaptive ways in and upon the physical world. Telephone books are paradigmatically data goods; specialized newsletters, being more selective, exemplify information goods; and brain surgery can be thought of as a knowledge good. We shall not elaborate further on these different types of good.

3. Information, Individuals, and Organizations

Most of what modern economics has to say about knowledge and information originates in the tradition of methodological individualism (Hodgson, 1988; 1993)[4]. This tradition takes the individual human being, *Homo Economicus*, as the fundamental unit of analysis. The origins of methodological individualism are deeply rooted in the Anglo-Saxon political economy tradition that goes back to Hobbes and to Locke (MacPherson, 1962). The central challenge was to protect the rationality postulate inherited from the Enlightenment from the centrifugal tendencies at work when varied and complex individuals pursue their own interests in both markets and organizations. The socialist calculation controversy of the 1930s opposed those who believed that rationality was best preserved through a central planning mechanism – a metamorphosis of Walras's auctioneer – to those who believed in preserving it through a decentralized market mechanism. The concern with the computational efficiency of either mechanism placed the focus on the coordinating role of ”knowledge” and ”information” (Von Mises, 1969; Lange and Taylor, 1938; Hayek, 1999) and on the computational characteristics of different types of economic agency – the state, the firm, the individual. More recent attempts to deal with these threats to economic rationality have resulted in a kind of methodological "cyborgism" (Mirowski, 2002) that builds information structures both above and between agents[5].

What are the computational requirements of the neoclassical rationality postulate? Most relevant from our perspective is the fact that, whatever the type of economic agent involved, it is not subject to communicative or data processing limitations. The information environment in which it operates is free of noise and friction – well-structured information is instantaneously available in the form of prices and these fully capture the relevant

[4] The Marxist tradition in economics has an even less tenable position on information than does neoclassical economics. In the Marxist tradition, information asymmetries are deliberately created for the purposes of exploitation. Information goods are or should be, by their nature, free goods. See Marx (1867).
[5] We are indebted to a reviewer of this paper for this observation.

attributes of trades into which the actor enters[6]. What does an individual economic agent actually do with information? He computes in order to take decisions (Hurwicz, 1969; Marschak, 1974; Arrow, 1973). His computational abilities are unbounded and it enjoys both infinite memory and infinite foresight (Stiglitz, 1983). It follows, therefore, that such an agent does not need to learn much (Hodgson, 1999). It is the frictionless ease with which the rational economic agent is able to compute and communicate that qualifies him as "Newtonian".

It is by now well established that *Homo Economicus* has not served neoclassical economics well as a model of the way in which real human beings process and transmit data. These agents are bounded in their rationality (Miller, 1956; Simon, 1945) and are subject to systematic cognitive biases (Kahneman and Tversky, 1982; Kahneman, 1994; Bruner, 1974; Jung, 1971). It was a coming to terms with the cognitive limitations of individual and group processes that gradually turned economics into what Mirowski calls a "Cyborg science", with Hayek as its prophet (Mirowski, 2002; Hayek, 1999; Gottinger, 1983; Makowski and Ostroy, 1993; Smith, 1991; Lavoie, 1985; Weimer and Palermo, 1982).

Evolutionary economics developed a more realistic – not to say "naturalistic" (Quine, 1969) - perspective on the role of knowledge in human affairs than has orthodox economics (Hodgson, 1993; Vromen, 1995; Hamilton, 1991; Nelson, 1994). The omniscience of agents is out! For Nelson and Winter (1982), for example, the routinization of firm activities is a response to information complexity. It is in rules and routines that a firm's knowledge is deemed to be stored. These then become the units of selection in an evolutionary process. Yet, as Fransman points out, the tight coupling of information and knowledge that is implied – with knowledge becoming little more than processed information - is unrealistic, since different agents may extract different knowledge from the same information (Fransman, 1998). Indeed, the variety on which evolutionary selection is effectively predicated, depends on it! Fransman himself goes on to associate information with data – a tight coupling in the other direction - and knowledge with belief.

If economists of different stripes have tended to conflate knowledge and information, sociologists, by contrast, have been more concerned with knowledge alone. Furthermore, sociology's point of departure is not the asocial atomized individual, but the embedded

[6] As Koopmans put it "The economies of information handling are secured free of charge (Koopmans, 1957).

socialized actor (Granovetter, 1985). Mead, for example, emphasized “the temporal and logical pre-existence of the social process to the self-conscious individual that arises in it” (Mead, 1934, p. 186). Thus, in contrast with the methodological individualism of economics, sociology “problematizes” the individual, often adopting a Vygotskian view that society should be the point of departure for looking at the evolution of human information processing capacities (Vygotsky, 1986). Durkheim and Mauss (1903), for example, analyzed primitive classification schemes as *collective* forms of representation. Sociology, then, typically starts with a multiple-actor perspective and gradually homes in on the single actor.

Finally, the sociology of knowledge tradition emphasizes the way in which power shapes collective representations (Mannheim, 1960; Habermas, 1987). By viewing human rationality as socially bounded by power and institutions (DiMaggio and Powell, 1983; Scott, 1989), sociology avoids the requirements for hyper-rationality that has plagued neoclassical economic thinking. Of course, since institutional economics borrows heavily from the sociology of institutions and organizations, issues both of bounded rationality and of power and influence have come to figure prominently in its analyses. They also figure in Agency theory and in theories of incomplete contracting (Jensen and Meckling, 1976; Hart, 1995; Grossman and Hart, 1988).

The new institutional economics aspires to bridge the gap between neoclassical economics and organization theory (Williamson, 1985; Furubotn and Richter, 1998). Yet it remains weighed down by the neoclassical perspective on information. It acknowledges the existence of friction in the transactional processes that drive the economic system, but offers little or no theorizing about it. At best, it can differentiate between markets - an external information environment in which data is well codified and can therefore flow freely - and hierarchies - an internal information environment in which the flow of data is viscous on account of the tacit nature of the knowledge involved. The first type of more analytically tractable environment has typically been the province of economists; the second, more qualitative environment has been left to organizational theorists.

Perhaps on account of its more qualitative material, organizational sociology has addressed the problem of knowledge in organizations, but not much that of data or information. Working in the interpretive tradition initiated by Weber, it has focused on sense-making, the process through which information is interpreted and converted by receivers into intelligible knowledge (Weick, 1995; Daft and Weick, 1984; Gioia and Chittipeddi, 1991). But how the codes on which information is borne come into being in the first place is a question that needs to be addressed before one can progress on to sense-making. Habermas's theory of communicative action, for example, sees meaning as something to be freely negotiated between interacting agents (Habermas, 1987). But can the idea of an open negotiation realistically apply to the codes that agents inherit and draw upon in their interactions? Such codes do much to shape the possible meanings that are up for negotiation. Some of the concepts that organizational sociologists apply to knowledge will also apply to information[7], but for this to yield a credible result, they would have to explore the nature of data as well as that of information.

4. The Contribution of Information Theory

The discipline that comes closest to doing this is information theory. But, originating as it does in an engineering tradition, information theory concerns itself primarily with the challenge of information *transmission* rather than with problems of information *content* or *meaning* (Nyquist, 1924; Hartley, 1928; Shannon, 1948). It is more abstract in its approach to information than is sociology, being concerned with the technical characteristics of communication channels independently of the nature of message sources, senders, receivers, or message destinations. It seeks to establish efficient encoding strategies for channels subject to noise.

By relating the definition of information to the freedom of choice we enjoy when we choose a particular message from among all the possible messages that we might transmit, it becomes possible to calculate the amount of information carried by that message. It turns out to be the inverse of its probability of appearance. Since within the framework provided by information theory, any message consists of a sequence of symbols drawn from a given repertoire of symbols, the theory allows one to assess the effectiveness of different coding schemes using different symbolic repertoires in a channel. Shannon's *Mathematical Theory of Communication* (Shannon and Weaver, 1949) yields a number of fundamental theorems which set theoretical limits to the amount of information that a channel of given capacity is

[7] Giddens's theory of structuration, for example, and his concepts of domination, signification and legitimation (Giddens, 1984) can be used to analyze the distribution of both knowledge *and* information in a social system, the nature and extent to which these are codified, and their normative status.

able to transmit, both in the presence and absence of noise. Whether or not the limit is reached in a particular situation will turn on the choice of symbolic repertoires and syntactic rules, as well as on the choice of coding scheme.

The amount of information that can be transmitted, then, is a function of the size of the available repertoire of distinct symbols or states that is available, the relationships between symbols – i.e., the syntax - as well as the degree of fidelity required given the amount of noise in the channel. Information theory is primarily concerned with maximizing the fidelity of transmission at an acceptable cost – Shannon and Weaver (1949) refer to this as a technical level problem (level 1). As Shannon took pains to point out in his 1948 paper, information theory is not particularly concerned with what the symbols actually *mean* - a semantic level problem (level 2) - or with whether a given message has the desired effect on a given message destination—an effectiveness level problem (level 3). These he viewed as problems to be addressed by social scientists rather than engineers. Shannon thus sought to offer a clear line of demarcation between information and knowledge.

Crucially, information theory takes the repertoire of symbols to be transmitted as *a given*. It does not ask how the repertoire came into being, whence the distinctness of the symbol system came from, or whether the symbolic repertoire was established by prior convention or through a gradual process of discovery. Yet, before we are in a position to extract information from a symbol, we first need to extract the information *that it is indeed a symbol* and hence an acceptable candidate for further processing. It must, therefore, be distinguished from other stimuli that might register with an agent as data In short, information theory ignores the question of data, of how a given repertoire of symbols – a pre-selected collection of states – gets itself registered with an agent as a *data set* from which information can then be extracted[8].

If, as we have argued, data is a discernible difference between two states, at a purely theoretical level, the limiting case of what constitutes a difference is given by the calculus. It defines, in the limit, what can ever count as data. Perhaps the *physically* limiting case of data is given by Planck's constant, which defines the smallest discernable event that can

[8] Interestingly, Blackwell applied the precepts of information theory to states rather than symbols. These could then acquire the status of commodities in an Arrow-Debreu analytical framework. Blackwell's work was to influence game-theoretic and other work on the economics of information (Blackwell and Girschik, 1954; Lipman, 1991; Plott and Sunder, 1978; Geanakoplos, 1992; Rubinstein, 1998).

pass off as a *state*. But for *us* as sentient beings, what counts as data is what we can actually discern. Our ability to discern differences between states only operates within a certain physiological range (Hargreaves Heap and Varoufakis, 1995). Outside that range, we cannot be sure that the different states that constitute our data are orthogonal to each other and hence capable of yielding a viable repertoire, as required by Shannon.

Data, then, and the regularities that reside within the data, are properties of events and things "out there" in the world – i.e., physical processes and products - that become available to us as sentient beings through our physiological apparatus, often amplified by instruments and other artefacts. Information, by contrast, is relational. As Bateson put it, it is "the difference that makes a difference"—and that means making a difference *to someone* (Bateson, 1971)**.** Thus we might say that regularities within data, an objective property of events and things, convey more or less information to different individuals, depending on their respective circumstances, such as their individual histories, their values and emotional make up (Damasio, 1999), their mental models, and the specific state of their expectations at any given moment.

The early founders of modern information theory—Nyquist, Hartley, Shannon—imported from thermodynamics the concept of entropy, which Shannon then associated with the amount of information *H* gained in a message. Building on the concept of entropy that information theory shares with thermodynamics, we would like to suggest that information-bearing data may be likened to free energy in a physical system. That is to say, data that carries information retains a capacity to do *work* – i.e., it can act on an agent's prior state of expectations and modify it. Data that carries no information may be likened to bound energy in physical systems: to the extent that it leaves an agent's state of expectations unmodified, it has performed no work on its expectational structure.

Note that we are dealing here with both an objective term—the quantity of information that can potentially be carried by a given data set[9] –and a subjective term - the amount of information that can be extracted in practice from the data set by a situated agent. When we claim that information is relational, it is with respect to the second of these terms. This "subjectivist" view of information, however, based as it is on an agent's *situated* expectations, confronts the "objectivist" view of information developed by Shannon, one

[9] This quantity has been calculated for different states of physical matter by Seth Lloyd (Lloyd, 2000).

that is based on *conventionalized* expectations. The English language, for example, contains an objective amount of information based on the relative frequency of appearance of letters of the alphabet and of certain types of words, such as articles and pronouns. In Shannon's view, information content is set by the ratio of actual to possible events. In the examples that he gives, however, both the repertoire of possible events and their frequency are fixed *a priori*, so that the computation of information content is straightforward. Yet, to stay with the example of the English language, as soon as we move up to the level of sentences and paragraphs, the number of possible sentence constructions moves to infinity. Does this mean that information content moves to infinity? No, simply that the repertoire of possible sentences is now largely shaped by the circumstances in which any given sentence is likely to be uttered – i.e., by its *context*. But *context varies in the extent to which it is shared across individuals*. Some contexts will be unique to individuals, while other contexts will be widely shared.[10]

Native English speakers, for example, will share almost identical expectations concerning the frequency of appearance of letters in English words. They will share somewhat similar expectations concerning the frequency of appearance of many classes of words in a well-constructed English sentence. They will share far fewer expectations, however, concerning the rate at which other words appear in the sentence, for these will depend on particular circumstances. The discourse that might take place in a biology laboratory, for example, will be meaningful to a much smaller group of people than the one taking place on a televized chat show. In sum, it is *shared context*, the generator of inter-subjective objectivity (Popper, 1959) that stops information content from ballooning to infinity and that renders discourse possible.

Shannon takes care of this difficulty largely by avoiding it. Given his focus, this was not unreasonable. As a communication engineer, he was concerned mainly with the objective and computable aspects of information and the requirements that these might impose on a communication channel. Thus Shannon addressed what he called the level 1 or technical problem (was the message received the same as the message sent?) and confined his analysis to well defined and delimitable repertoires. What he called the level 2 or semantic

[10] Information in the objectivist view can be seen as the higher bound of the ensemble of all possible “subjectivist” or “inter-subjectivist” interpretations that could be extracted from the data. Yet in any but the most simple contexts, the objectivist view confronts a Godelian ‘undecidability’ problem .

problem (is the received message understood?) was not his concern. This depended on whether the receiver possessed the relevant code – i.e., some familiarity with the alphabet, the vocabulary and the syntactic rules of the English language, etc. Note that, even here, the repertoire was assumed by Shannon to be closed: the alphabet is limited in size as are both the vocabulary and the syntactic rules that have to be attended to. Finally, what Shannon called the level 3 or effectiveness problem (does the message lead to the desired behaviour?), was completely outside his purview. Both levels 2 and 3 we identify with knowledge.

It is clear that, where symbolic repertoires and syntactic structures are established by convention rather than by discovery, technical level (level 1) communication issues need not concern themselves with the idiosyncratic characteristics of communicating agents. However, the minute we move to the semantic level (level 2) or to the effectiveness level (level 3), the dispositional states of the agents - i.e., their prior knowledge - become relevant. Agents are *situated* processors and transmitters of data. The individual agent's *memories* as well as his preference orderings - and hence *values* and *emotional dispositions* (Damasio, 1999) - therefore need to be reckoned with. It is at levels 2 and 3, then, that the idiosyncrasy and potential subjectivity of context becomes most manifest. Here, selection is constrained less by rules than by personal style and preference.

We can represent the issue that we are discussing with a diagram In the rectangle of figure 2, we variously mix expectations—and hence probabilities—based on agreed conventions concerning what constitutes an event, the number of recurrences of that event that constitute a fair sample, etc., with expectations based on personal experience. The first type of probability will lend itself to a frequency interpretation whereas the second will lend itself to a Bayesian or subjectivist interpretation. We subdivide the rectangle into three zones and associate each zone with one of Shannon's three levels. We see that Shannon's level 1 problem—the technical problem—leaves little or no scope for the subjectivist approach to probability. It is also the level that is the most computationally tractable and the one to which Shannon himself decided to confine his analysis. His level 2 problem—the semantic problem—is one that offers somewhat more scope for subjective probabilities to kick in. In language, for example, syntactic constraints and word usage will conventionalize expectations to some extent, but personal idiosyncrasies and style will inject a strong subjective element into the communication process. Finally, Shannon's level 3 problem—

the pragmatic problem—leaves little scope for the frequency perspective, since at this level, conventions hardly appear as anything other than subjectively experienced and highly variable constraints.

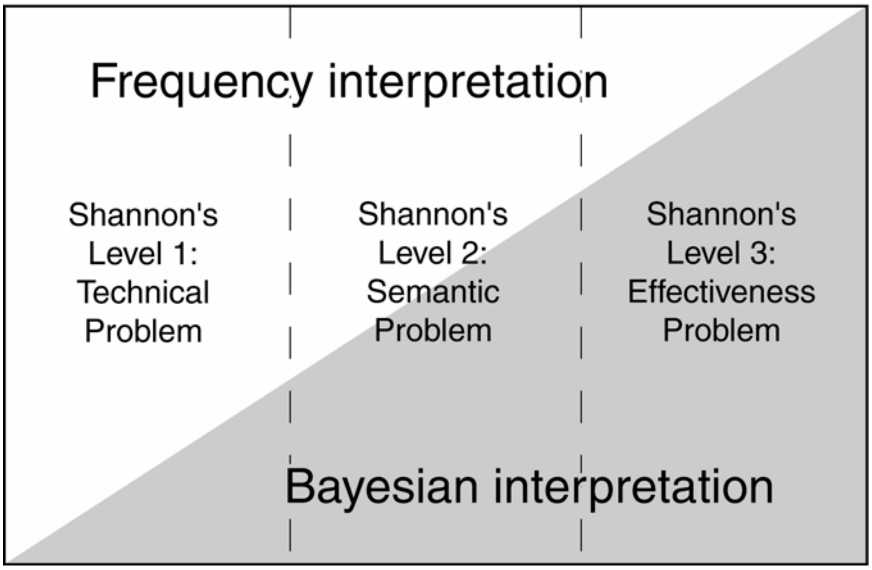

Fig. 2. Frequency and Bayesian interpretations in Communication

The implication of the above is that, whatever intrinsic regularities it contains, to the extent that data only carries information when it can modify an expectation, what constitutes data *tout court* for me, might turn out to be information-bearing data for you. How far we are aligned in our information-extraction strategies will depend on how far our respective expectations are shaped by conventions, that is, socially shared encoding rules and contextualizing procedures, or by idiosyncratic circumstances - codes and contexts that are not widely shared. The act of extracting information from data constitutes an *interpretation* of the data. It involves an assignment of the data to existing categories according to some set of pre-established schemas or models that shape expectations. For this to be possible, *such schema or models must already exist in some form or other.*

But how do such schemas and models come into existence in the first place? They do so primarily through explicit or tacit rules of inference. Explicit rules will for the most part be applied to codes; implicit rules will be applied primarily to context. Expectations and categories co-evolve, with expectations shaping the categories that we create, and these, once created, in turn shape the evolution of subsequent expectations[11]. Our categories condition the dispositions that we adopt towards the world – i.e., our knowledge, taken here in the Popperian sense of a disposition towards action (Popper, 1983)[12]. *Thus, data can only constitute information for an agent who is already knowledgeable***.** Data can be viewed as a low energy system that acts informationally rather than mechanically (Boisot, 1995), that is to say, it gives rise to *intentional* action rather than mere mechanical movement[13]. Guided by the structure of its expectations, an agent first extracts what constitutes information for him from the regularities available in a data stream and then acts upon it (see Figure 1).

Given its almost exclusive focus on the technical level of communication, the work of information theory has largely ignored such issues. These, occurring as they do at Shannon and Weaver's level 2 and their level 3 – i.e., at the level that we have identified with knowledge rather than information - have proved to be of more interest to interpretative sociology. In organizational sociology, the semantic problem shows up as a concern with sense-making (Weick, 1995) or bounded rationality (Simon, 1945; Kahneman and Tversky, 82), whereas the pragmatic problem shows up as a concern with power, values, and influence (Habermas, 1987). Level 2 and level 3 problems are also of relevance to institutional theory (DiMaggio and Powell, 1983; Scott, 1989). It is clear, however, that at these levels, we are far removed from an economic world in which agents can be assumed to have common knowledge of rationality, a consistent alignment of beliefs, and rational expectations (Aumann, 1976). But if information theory's concern with bits and bytes led it to shun the issue of knowledge, it also managed to sidestep the issue of data by assumption: for all intents and purposes, information and data were the same thing. Although it never explicitly claimed otherwise, almost unwittingly the new discipline of information physics has highlighted the issue.

[11] Kantians believe that the categories came first, while Lockeans believe that expectations came first and that these were shaped inductively by the recurrent features of our experiences. The debate continues; mercifully, we need not get involved.
[12] Kenneth Arrow has the same expectational view of information as Popper does. See Arrow (84).
[13] This is not to say that both informational and mechanical effects cannot be present at the same time. But where energy acts informationally, we can affort to ignore its mechanical effects on behaviour. These are negligible.

5. The Physics of Information

According to the late Rolf Landauer, "Information is physical" (Landauer, 1999), and the most fundamental analysis of the nature of information so far carried out originates in physics. Even within physics itself, since the most fundamental analysis of physical processes takes place at the quantum level, it is within the new field of quantum information theory that we confront the deepest level of analysis of information An important breakthrough for the development of quantum information theory was the discovery that quantum states could be treated *as if they were information* (Nielsen and Chuang, 2000). Thus, if information is physical, *what is physical is also information* (Lloyd, 2000). Quantum information theory, being broader in scope than classical information theory, operates at the most abstract level, quite removed from any social science conception of information. Can such a view of information have anything to offer the social sciences?

If information is physical, then, like any other physical process, it is subject to the second law of thermodynamics. The physical entropy involved here, however, must be distinguished from the Shannon entropy, even though the two are closely related. One might, in effect, say that Shannon entropy is predicated upon thermodynamic entropy. In a closed system, both types of entropy-generating processes turn out to be irreversible[14].

Although physicists have not much concerned themselves with it, the distinction that we are drawing between data, information and knowledge is implicit in the work being done in the Physics of Information (Zurek, 1990; Feynman, 1996; Feynman, 1999; Bennett, 1999; Landauer, 1999). If the bit is the fundamental unit of analysis in classical information theory, then the qubit is the fundamental unit of analysis in quantum information theory. Just as a classical bit is in one of two possible states, 0 or 1, so a qubit has two possible *eigenstates* $|0\rangle$ or $|1\rangle$. One difference between a bit and a qubit, however, is that the latter can also be in any well-defined linear combination of the two eigenstates, $|0\rangle$ or $|1\rangle$. Another difference is that, whereas we can directly examine a bit to determine what state it is in, we cannot directly examine a qubit to determine its quantum state without destroying that state. In short, *the eigenstates of the qubit are not available to us as data.*

[14] It turns out that, at the quantum level, not all computations are irreversible (Bennett, 1999).

By the postulates of quantum mechanics, any measure that we perform on a qubit reduces it to one of its eigenstates. This dichotomy between the state of the qubit and what we can observe lies at the heart of quantum information and quantum computation. At the quantum scale, we can ask: how much information does a qubit represent in the absence of measurement? It turns out that nature holds a great deal of "hidden information" in this way, and it grows exponentially with the number of qubits[15].

In the classical world, we assume that we can distinguish, at least in principle, between different states, since this is what qualifies them as data. Yet in the quantum world, we have to abandon such an assumption, for, unless the orthogonality between two given states can be maintained, one can no longer readily distinguish between them and register such a distinction as data. Without data one cannot extract reliable information from the system concerning such states. It turns out that, below a certain scale known as the Planck scale, the orthogonality between two states can no longer be securely established. *There are thus physical limits to our access to data and hence to our ability reliably to extract information from data.*[16]

These limits first appeared in 1867 in the field of thermodynamics in the shape of Maxwell's Demon, a microscopic creature that appeared to violate the second law by using information to distinguish between fast- and slow-moving particles and hence to throw dissipative processes into reverse (Leff and Rex, 1990). To understand the nature of the thermodynamic limits on our access to data, we can revert to our earlier and perhaps somewhat oversimplifying analogy, taking data as corresponding to the general category of energy, and information-bearing data as corresponding to free energy - i.e., it has a capacity to do work in the sense that it can modify our expectations, and, hence, the state of our knowledge. Noise would then correspond to bound energy: it either consists of data that carries no information for us and can therefore do no work, or it consists of states that cannot be distinguished from one another and that hence do not even graduate to the status

[15] Some, notably Penrose, have argued that quantum effects are also manifest in human cognitive processes (Penrose, 1994; Green, 2000). We cannot, however, observe each other's mental states directly without disturbing these states; we can only observe the behavioural outputs of these states.
[16] That there are *biological* limits to our access to data as well has been know since the work of Fechner in the nineteenth century.

of data for us. Noise cannot modify our expectations. Like bound energy, it can perform no work.

Now, although knowledge itself is dispositional, it reveals itself in purposeful agent behaviors such as data processing, data transmission and actions based on these. We hypothesize that data storage, on the one hand, and data processing, transmission and purposeful action, on the other, can be respectively likened to the build up of potential energy and to its subsequent exploitation as kinetic energy. As a disposition to act, then, knowledge corresponds to potential energy - a stock; and as purposeful action or behavior, knowledge corresponds to kinetic energy - a flow[17]. In open systems, both the transformation of potential energy into kinetic energy, and the transformation of the latter into work, are subject to dissipation.

Landauer (1990) demonstrated that energy dissipation occurs both in the information storage process as well as in the information transmission process as a result of *information erasure*. No information is erased in a reversible computation, however, because the input can always be recovered from the output. When we say, therefore, that a computation is reversible, we are really saying that no information is erased during the computation. *Landauer's principle* provides the link between energy dissipation and irreversibility in computation, stating that, in order to erase information, it is necessary to dissipate energy. The principle can be stated thus (Landauer, 1990):

> If a computer erases a single bit of information, then the amount of energy dissipated into the environment will be *at least* $k_bT \ln 2$, where k_b is a universal constant known as *Boltzmann's constant*, and T is the environmental temperature of the computer.

The laws of thermodynamics also allow us to express Landauer's principle in terms of entropy rather than in terms of energy and dissipation (Landauer, 1990):

> If a computer erases a single bit of information, then the environmental entropy will increase by at least *at least* $k_b \ln 2$, where k_b is Boltzmann's constant .

[17] There are, of course, important differences between knowledge and potential energy, on the one hand, and behaviour and kinetic energy, on the other. A stock of knowledge is not depleted in the way that a stock of potential energy might be. It can only be *dissipated*, in the sense that the information structures that constitute the stock are gradually eroded. Likewise, behaviors, through the mechanism of *learning* can actually build up a knowledge stock rather than depleting it - which is what kinetic energy does to potential energy. Clearly, reasoning by analogy must know its limits.

It should be noted that Landauer's principle effectively provides us only with a lower bound on the amount of energy that must be dissipated to erase information. Clearly, if all computational processes were reversible, then the principle would imply no lower bound on the amount of energy dissipated, since no bits would in fact be dissipated during computation (Nielsen and Chuang, 2000).

Maxwell's Demon is located at the meeting point of a physics of energy and a physics of information. Bolzmann's definition of entropy links the two types of physics. An informational limit is reached under two quite different conditions. The first occurs when the energy expenditures incurred by data capture and transmission activities required to distinguish between two states, and hence to create discernable data - often performed by specialized equipment - itself exerts a mechanical effect on those states, thus preventing them from stabilizing enough to get themselves detected - Heisenberg's uncertainly principle describes this condition. The mechanical effects of such energy expenditures then swamp and overwhelm their informational effects. The second occurs when the Demon needs to store transmitted data in memory for subsequent processing. Assuming that the Demon's memory is finite - i.e., it is subject to bounded rationality (Simon, 1945) - it will sooner or later confront the need to erase stored data in order to make way for new data. Landauer's principle tells us that at that moment data will be lost and entropy levels - both thermodynamic and informational - will increase. However, thermodynamic entropy and information entropy are quite distinct from one another. Although both draw on Boltzmann's formula, the first refers to the regularities or lack of them in discernible states-of-the-world – that is, in data - whereas the second refers to the information that can be extracted from such states by a knowledgeable observer. In sum, if social scientists conflate information and knowledge, physicists conflate data and information. In the next section, by means of a simple diagram, we will indicate why both types of conflation matter.

6. An economic interpretation of the principle of least action

Any physical system is subject to the principle of *least action*, an integral variational principle initially put forward by Maupertuis in 1744 that establishes the difference between the actual motion of the system and all of its kinematically possible motions during a finite time interval (Barrow and Tipler, 1986). According to Green, when the observables of the system, such as its energy, its momentum, its angular momentum, its central vector and certain other charges, have prescribed values on the boundary of any region of space and time, they will vary in such a way that the total action within the region has its minimum value (Green, 2000). This will be as true of dissipative systems as it will be of Hamiltonian systems. To the extent that the system has a capacity for storing memories of earlier states - and this does not require that the system be intelligent or even alive - then it will be able to use data and information in such a way as to minimize the action.

Being universal in scope (Omnes, 1999), the principle of least action implies that nature as a whole makes choices that are economic in their outcomes[18]. How might Maxwell's Demon apply the principle? In effect, it allows us to posit the existence of a trade-off between the Demon's consumption of energy and his consumption of data resources as it attempts to sort out fast-moving from slow-moving particles. Such a trade-off can usefully be represented by means of a scheme that is somewhat reminiscent of a production function, but the purpose of which is limited to illustrating the economic nature of the principle of least action.

A production function is a schedule showing the maximum amount of output that can be produced from any specified set of inputs (Fergusson, 1969). In neoclassical production functions that take capital and labor as inputs, information and knowledge are not explicitly represented as factors of production in their own right, although, in talking of capital and labor, we may take them to be implicitly present. The knowledge embedded in machinery and equipment, for instance, clearly forms part of the capital factor, and the labor factor clearly embodies the know-how and experience of employees. Given that in the so-called “new economy” information and knowledge have clearly moved center-stage, some have claimed that they should therefore become critical productive factors in their own right alongside capital and labor (Bell, 1973; Romer, 1986; Romer, 1990)[19].

Yet, given that information and knowledge are already implicitly embedded in traditional productive factors, this would result in double counting. As an alternative, therefore, one

[18] In the nineteenth century, this was referred to as *the economy of nature*.
[19] By the term new economy, we mean more than an economy driven by the Internet phenomenon. We therefore avoid having to take sides in the current debate as to whether there is in fact a new economy.

could move up to a more abstract and general level and bring together two different classes of productive factors: 1) purely physical factors, such as space, time, and energy – these would be measured in physical units such as meters, seconds, and joules, and 2) data factors, being discernible differences in the states of the physical factors – these would be measured in bits (Boisot, 1995; Boisot, 1998). Note that, in this new scheme, information and knowledge are not taken as factors of production at all. According to our earlier arguments, information constitutes an extraction from the data factor that results in *economizing* on that factor and hence in a move towards the origin. Knowledge, likewise, economizes on data-processing - and hence on the consumption of data inputs - more so in the case of abstract knowledge than of concrete knowledge[20].

An example of the new scheme is shown in figure 3. As will shortly be apparent, much as it may look like one, it is not actually a production function. In the diagram, we can distinguish two types of movement, one *along* isoquants and another *across* them. A move to the left along an isoquant represents a progressive substitution of data for physical factors, something that happens when, by gradually accumulating the data of experience, systems "learn-by-doing", with less expenditure of time, space, and energy, in whatever task they are performing – manufacturing aircraft wings, miniaturizing electronic components, etc. Learning-by-doing can only work for systems that can store past states – i.e., for systems that have *memory*. Some purely physical systems have memory and all living systems do. By implication, a move to the right along an isoquant can be interpreted either as forgetting, an erosion of memory, or as the workings of bounded rationality. Both rightward and leftward movements are possible. A downward vertical movement across isoquants and towards the origin in the diagram, by contrast, represents the generation of *insight*, the extraction of information from data to create new, more abstract knowledge concerning the structure underlying phenomena. This second movement – the joint effects of pattern recognition and computational activities – is discontinuous, reflecting the unpredictable nature of creative insights (Miller, 1996). It makes it possible to reach the same output levels as before with less data processing and hence a lower consumption of data inputs. In addition to having memory, a system that has a capacity for insight must also be *intelligent*, that is, it must be capable of processing data in order to extract information from it in the form of patterns or structures.

[20] Ernst Mach's "Principles of the economy of thought" were an important source of inspiration for Hayek (Mirowski, 2002).

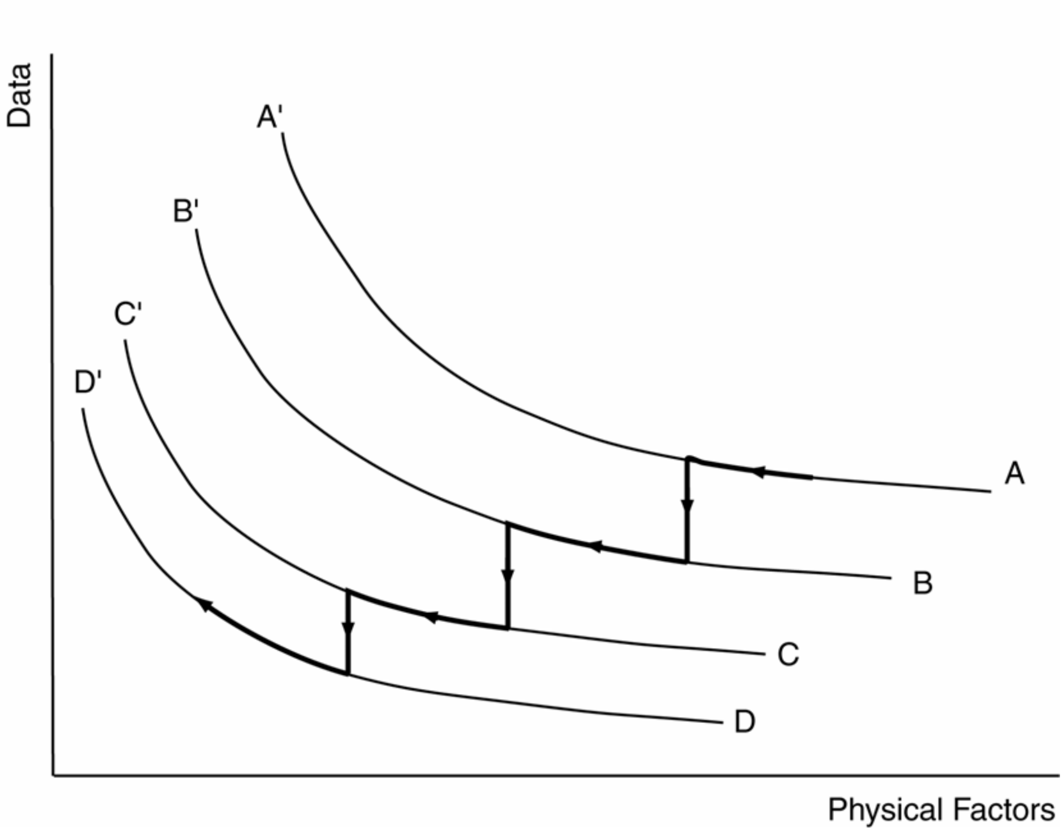

Fig. 3. Data vs. physical factors scheme

Our scheme and the neoclassical production function have some similarities. For example, they both take movement along an isoquant as representing *technical change* – i.e., a change in the mix of data and physical resources that generate a given output - and movement across isoquants towards the origin as representing *technical progress* (Boisot, 1998) – i.e., a reduction in the quantity of data and/or physical resources required to generate that output. Yet the two schemes differ in three important ways.

First, while the neoclassical production function offers no preferred direction for movements along an isoquant, the broad tendency to substitute data factors for physical factors in our scheme – a process of variation, selection and retention that results in data accumulating in the form of memory — imparts a direction to technical change and to technical progress. Why? The clue resides not in the evolutionary nature of knowledge – although this is certainly a factor – but in the evolutionary nature of agents, individual or corporate (Metcalfe, 1998). To the extent that evolution enhances both the memory and the data processing capacity of agents – this can be achieved either via biological evolution or via the artefacts of cultural evolution (Boyd and Richerson, 1985; Clark, 1997) - they are able to make better use of whatever data accumulates over time, and this at a lower cost than that of using the physical resources available to them. Thus, if intelligence is selected for by evolution, intelligence, in turn, will demonstrate a selection bias in favor of data over physical resources

In effect, in contrast to the neoclassical production function in which movement along an isoquant is reversible, the arrow of time is at work in our scheme, allowing it to describe irreversible and hence path-dependent processes that, according to circumstances, might be characterized as being either evolutionary or as developmental. We must emphasize that the arrow of time manifests itself in global rather than local behaviors. The general tendency for a leftward movement up an isoquant is likely to have many local exceptions - brought about either by forgetting or by bounded rationality – that move it in the opposite direction.

Second, our scheme is able to account for technical progress. Although in both the neoclassical and in our scheme, technical progress in described by a jump across isoquants towards the origin, in the neoclassical case, such a discontinuity cannot be explained; it had to be exogenously given. In our scheme, by contrast, a discontinuous jump from one isoquant to another is accounted for by a discontinuous jump in a living system’s own learning processes – i.e., it is accounted for by the discontinuous phenomenon of *insight*, the extraction of informative patterns or gestalts from data, and their subsequent conversion into knowledge.

Third, the data and physical factors that make up our scheme present quite distinct economic properties. While, in the neoclassical production function, purely physical factors are naturally subject to scarcity and hence appropriable, the data factors of our own scheme are not. While they may not always be immediately accessible—and in that sense they may be considered scarce - once one has secured them, they can often be replicated and distributed at almost zero marginal cost. Providing it has found the right physical substrate, therefore, data will propagate rapidly and extensively. Scarcities will then only appear in the form of a living system's limited capacity to receive, store, process, and transmit data, not in the data factors themselves[21]. For this reason, data factors are much more difficult to appropriate and to subject to traditional forms of economic exchange than purely physical factors. They are hard to price and this makes it hard to use price signals to guide a substitution of data factors for physical ones. Our scheme can illustrate such a substitution process; it cannot analyze it.

To summarize: except, perhaps, for the universe as a whole, there are no perfectly closed systems in the real world. Open systems are prey to unwanted interactions with their environment that get registered as noise when viewed informationally. Economics has tended to ignore the implications of the fundamental openness of the systems they study. Our scheme, by allowing the representation of the effects of time and entropy in the economic process, rectifies the situation. Once you admit learning, development and evolution into the picture, you admit irreversible processes. But our scheme also suggests that the entropy concept is but one side of the coin when dealing with the second law of thermodynamics. Irreversible processes can lead to emergent, order-creating outcomes as well as to entropic ones, those that allow living things to jump across isoquants and move towards the origin in pursuit of factor savings (Brooks and Wiley, 1988). Although we may agree with Shapiro and Varian when they observe in *Information Rules* (1999) that the information economy has not yet repealed the laws of economics, we feel that it poses explanatory challenges to economics—well captured by the way that novelty and new knowledge emerges in living systems and organizations—that the discipline has yet to take on board.

7. Implications

We can briefly summarize our discussion in the following three propositions:

[21] We can see the logic of our new scheme at work in the way that organizations are today attempting to handle large amounts of transactional data. Data mining, for example, is the process of extracting information from data. People will not pay for data, but as information extraction becomes ever more difficult and user-specific - i.e., customized - people will pay for information. To the extent that data can be turned into information that has relevance for someone, that someone will in principle be willing to pay for the data processing and transmission economies on offer. What such economies offer is the possibility of reallocating a key data processing resource possessed by all intelligent agents in finite quantities: *attention*. An information economy, by implication, has to be an attention economy as well.

1.Information is physical (Landauer, 1999). It is a constituent element of all physical processes and hence cannot treated as something epiphenomenal to the economic process. It must be engaged in on its own terms.

2.Economic agents subject to the principles of least action and to the effects of the second law of thermodynamics aim to economize on their consumption of both physical *and* data resources by deploying effective cognitive and behavioral strategies.

3.Effective cognitive strategies extract information from data and then convert it into knowledge. Effective cognitive and behavioral strategies vary from agent to agent as a function of their situation, of their prior individual knowledge, of their values, and of their emotional dispositions.

What follows from our three propositions?

Developing further the difference between data, information and knowledge, data generates *thermodynamic* entropy, which we shall label *entropy 1*. It involves the erasure of differences between physical *states*. Information, by contrast, generates *Shannon* entropy, which we shall label *entropy 2*. It involves the erasure of differences between *symbols*. The difference between physical states might well be maintained, but the *form* given to such states no longer yield unambiguous symbols. Finally, knowledge generates *cognitive* entropy, which we shall label *entropy 3*. It involves the erasure of differences between the possible *contexts* required for the interpretation of either states or symbols.

All these different types of entropy constitute variations on Bolzmann's formula

$$N \sum p_i \log p_i. \qquad \text{for } i = 1 \ldots n,$$

where n describes either the number of possible data states, the number of symbols in a repertoire, or the number of interpretative contexts that are compatible with a given set of states or symbols. N gives the message length, and we hypothesize that an inverse relationship exists between N and n. Efficient coding, however, should reduce both N and n to the extent that is builds on correlations between states, symbols, or interpretative contexts. Where such correlations are not given *a priori*, they must be discovered. In the absence of memory, however, an agent has no way of discovering such correlations so that, in effect, *n* can now potentially increase without limit. This makes Boltzmann's formula meaningless since it cannot be used as a basis for stable expectations.

Entropies 1 and 2 are to be found at Shannon's technical level. Entropy 3 is to be found at Shannon's semantic and effectiveness levels. At the semantic level, it can occur because the receiver does not know the codes or what, *specifically*, they refer to—this, in effect, is context narrowly defined—and at the pragmatic level it can occur because the receiver does not know to embed the *message as a whole* into an appropriate context. Entropy 1 has the effect of increasing Entropy 2 and Entropy 3. However, redundancy at the semantic and effectiveness levels can mitigate the effects of entropy 1.

Economics at best has only ever operated at Shannon's technical level. By largely ignoring problems of meaning and values, it has only scratched the surface of Shannon's semantic and effectiveness levels. Yet the implication of our analysis is that, strictly speaking, *there is no such thing as common knowledge and there is common information only to a limited extent. Only data can ever be completely common between agents.* As Metcalfe puts it, agents may *live* in the same world, but they *see* different worlds (Metcalfe, 1998). In its treatment of information, economics thus fell between two stools. On the one hand, it eschewed the complexities of the "soft" approach to knowledge and information associated with the semantic and effectiveness levels - and with the social and cognitive sciences as a whole. On the other hand, it never really dug into the foundations the way that physics did, in order to distinguish entropies 1 and 2 from one another[22]. It therefore allowed the concept of information to take whatever form was needed to maintain analytical and computationally convenience.

8. Conclusion

From "soft" sciences such as sociology right across to "hard" sciences such as physics, information has become a central concern. Economics, however, has tended to treat information as something unproblematic, an auxiliary concept that can be left largely

[22] Yet if the physics of information helped effectively, to distinguish entropy 1 from entropy 2, this did not result in a distinction within the discipline between data and information.

unanalyzed. Yet, in post-industrial economies, information has now become the main focus of economic transactions, and not merely a support for them. Economists, therefore, cannot afford the luxury of neglecting the conceptual foundations of an economics of information in this way.

Shapiro and Varian have argued that the laws of economics apply to the information economy no less than to the energy economy that preceded it. This is undoubtedly true and certainly needed to be said. The issue, however, is not about the *applicability* of economic laws, but about their *scope*. The physics of Newton was not displaced by that of Planck and Einstein, rather it ended up having to share the stage with them. Likewise, the physics of information neither falsifies the economic laws of the energy economy, nor does it render them irrelevant. What it brings out, however, is that, given their limited engagement with the concept of information, such laws will have trouble dealing with many of the phenomena associated with the evolution and growth of knowledge in general and with the emergence of the new economy in particular. They therefore need to be complemented with more encompassing and general laws that take into account the pervasive roles played, respectively, by data and information in all physical processes, as well as that played by knowledge in biological ones. Such roles are distinct and complementary and in need of clear articulation. This paper has attempted to provide some initial theoretical reflections on a task that still lies ahead.

References

Arrow K J (1984) The Economics of Information. Belknap Press of Harvard University Press, Cambridge, MA.

Arrow K J (1984) On the agenda of organizations. In: Arrow K J (ed) The economics of information 1st edn. pp 167-184. Belknap Press of Harvard University Press, Cambridge, MA.

Arrow K J (1996) The Economics of Information: An Exposition. Empirica 23 : 119-28

Aumann R J (1976) Agreeing to disagree. Annals of Statistics 4 : 1236-1239

Barrow J D, Tipler F J (1986) The Anthropic Cosmological Principle. Oxford University Press, Oxford

Bateson G (1971) The Cybernetics of 'Self': A Theory of Alcoholism. Psychiatry 34 (1): 1-18

Bell D (1999) The Coming of Post-Industrial Society. Basic Books, New York

Bennet C H (1999) Quantum Information Theory . In: Hey A J G (ed) Feynman and Computation: Exploring the Limits of Computers pp 177-190. Perseus, Cambridge, MA.

Binmore K (1990) Essays on the Foundations of Game Theory. Basil Blackwell, Oxford

Blackwell D, Girschik M A (1954) Theory of Games and Statistical Decisions. Wiley, New York

Boisot M H (1995) Information Space: A Framework for Learning in Organizations, Institutions and Culture. Routledge, London

Boisot M H (1998) Knowledge Assets: Securing Competitive Advantage in the Information Economy. Oxford University Press, New York

Boyd R, Richerson P J (1985) Culture and the Evolutionary Process. The University of Chicago Press, Chicago

Boyle J (1996) Shamans, Software and Spleens: Law and the Construction of the Information Society. Harvard University Press, Cambridge, MA.

Brooks, D., and Wiley, E., (1986) Evolution as Entropy: Toward a Unified Theory of Biology, Chicago: The University of Chicago Press

Bruner J S (1974) Beyond the Information Given: Studies in the Psychology of Knowing. George Allen and Unwin, London, UK.

Clark A (1997) Being There: Putting Brain, Body, and World Together Again. The MIT Press, Cambridge, MA.

Crary J (1999) Suspensions of Perception: Attention, Spectacle, and Modern Culture. The MIT Press, Cambridge, MA.

Daft R, Weick K E (1984) Towards a Model of Organizations as Interpretation Systems. Academy of Management Review 9 (2): 284

Damasio A R (1999) The Feeling of What Happens: Body and Emotion in the Making of Consciousness. Hartcourt, San Diego

Deleuze G (1969) Logique Du Sens. Les Éditions de Minuit, Paris

Derrida J (1967) L'Écriture Et La Différence. Les Éditions du Seuil, Paris

DiMaggio B, Powell W W (1983) The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. American Sociological Review 82 : 147-160

Durkheim E, Mauss M (1963) Primitive Classification. Cohen and West, London, UK.

Fergusson, C.E., (1969) Microeconomic Theory, Homewood, Illinois: Richard D. Irwin

Feynman R P (1996) Feynman Lectures on Computation. Addison-Wesley, Reading, MA.

Feynman R P (1999) There's Plenty of Room at the Bottom. In: Hey A J G (ed) Feynman and Computation: Exploring the Limits of Computers pp 63-76. Perseus, Cambridge, MA.

Fransman M (1998) Information, Knowledge, Vision and Theories of the Firm. In: Dosi G, Teece D J,Chytry J (eds) Technology, Organization, and Competitiveness: Perspectives on Industrial and Corporate Change pp 147-191. Oxford University Press, New York

Furubotn E, Richter R (1998) Institutions and Economic Theory: The Contribution of the New Institutional Economics. The University of Michigan Press, Ann Arbor, MI.

Geanakoplos J (1992) Common Knowledge. Journal of Economic Perspectives (6): 53-82

Giddens A (1984) The Constitution of Society: Outline of the Theory of Structuration. Polity Press, Cambridge

Gioia D, Chittipeddi K (1991) Sensemaking and Sensegiving in Strategic Change Initiatives. Strategic Management Journal 12 : 433-448

Gottinger H-W (1983) Coping With Complexity. Reidel, Boston, MA.

Granovetter M (1985) Economic Action and Social Structure: the Problem of Embeddedness. American Journal of Sociology 91 : 481-501

Green H S (2000) Information Theory and Quantum Physics: Physical Foundations for Understanding the Conscious Process. Springer, Berlin

Grossman S, Hart O (1988) One Share-One Vote and the Market for Corporate Control. Journal of Financial Economics 20 : 175-202

Habermas J (1987) The Theory of Communicative Action I: The Critique of Functionalist Reason. Polity Press, Cambridge

Hamilton D B (1991) Evolutionary Economics: A Study of Change in Economic Thought. Transaction Publishers, New Brunswick

Hargreaves Heap S, Varoufakis Y (1995) Game Theory: A Critical Introduction. Routledge, London, UK.

Hart O (1995) Firms, Contracts, and Financial Structure. Oxford University Press, New York, NY.

Hartley R V L (1928) Transmission of information. Bell System Technical Journal : 535-563

Hayek F A (1999) The use of knowledge in society. In: Cortada J W, Woods J A (eds) The knowledge management yearbook 1999-2000 1st edn. Butterworth-Heinemann, Woburn, MA

Hirshleifer J, Riley J J (1992) The Analytics of Uncertainty and Information. Cambridge

University Press, Cambridge, UK

Hodgson G (1988) Economics and Institutions: A Manifesto for a Modern Institutional Economics. Polity Press, Cambridge

Hodgson G (1993) Economics and Evolution: Bringing Life Back into Economics. Polity Press, Cambridge

Hodgson G (1999) Economics and Utopia: Why the Learning Economy Is Not the End of the History. Routledge, London

Hurwicz L (1969) On the Concept and Possibility of Informational Decentralization. American Economic Review 59 : 513-34

Jensen M, Meckling W (1976) Journal of Financial Economics 3 : 305-60

Jung C G (1971) Collected Works: Vol. 6, Psychological Types. Princeton University Press, Princeton, N.J.

Kahneman D (1994) New Challenges to the Rational Assumption. Journal of Institutional and Theoretical Economics 150 : 18-36

Kahneman D,Tversky A (1982) Judgment under uncertainty: Heuristic and biases. In: Kahneman D, Slovic P,Tversky A (eds) Judgment under uncertainty: Heuristic and biases pp 3-20. Cambridge University Press, Cambridge, UK.

Koopmans T (1957) Three Essays on the State of Economic Science. Kelley, New York

Kuhn, T.S, (1974) 'Second Thoughts on Paradigms', in Suppe, F., (ed), The Structure of Scientific Theories, Urbana: The University of Illinois Press

Kuhn T S (1996) The Structure of Scientific Revolutions. 3a edn.: The University of Chicago Press, Chicago

Lamberton D M (1996) The Economics of Communication and Information. Elgar, Cheltenham

Lamberton D M (1998) Information Economics Research: Points of Departure. Information Economics and Policy 10 : 325-330

Landauer R (1990) Irreversibility and heat generation in the computing process. In: Leff H S, Rex A F (eds) Maxwell's Demon: Entropy, Information, Computing pp 188-196. Adam Hilger, Bristol, UK.
Notes: Originally in IBM J. Res. Dev., 5, 183-91 (1961)

Landauer R (1999) Information is Inevitably Physical. In: Hey A J G (ed) Feynman and Computation: Exploring the Limits of Computers pp 77-92. Perseus, Cambridge, MA.

Lange O, Taylor F M (1938) On the Economic Theory of Socialism. University of Minnesota Press, Minnesota

Lavoie D (1985) Rivalry and Central Planning. Cambridge University Press, Cambridge, UK.

Leff H S, Rex A F (1990) Maxwell's Demon: Entropy, Information, Computing. Adam Hilger, Bristol, UK.

Lipman B L (1991) How to Decide How to Decide How to... Econometrica 59 : 1105-25

Lloyd S (2000) Ultimate physical limits to computation. Nature 406 (31): 1047-54

MacPherson C (1962) The Political Theory of Possessive Individualism: Hobbes to Locke. Oxford University Press, Oxford, UK.

Makowski L,Ostroy J (1993) General Equilibrium and Market Socialism. In: Bardhan P, Roemer J (eds) Market Socialism: The Current Debate Oxford University Press, New York

Mannheim K (1960) Ideology and Utopia. Routledge and Kegan Paul, London, UK.

Marschak J (1974) Economic Information, Decision and Prediction: Selected Essays, 3 Vols. Reidel, Dortrecht

Marx K (1970) Capital: A Critique of Political Economy. Lawrence and Wishart, London

Mead G H (1962) Mind, Self and Society: From the Standpoint of a Social Behaviourist. The University of Chicago Press, Chicago, IL.

Metcalfe J S (1998) Evolutionary Economics and Creative Destruction. Routledge, London, U.K.

Miller, A,I, (1996) Insights of Genius: Imagery and Creativity in Science and Art, Cambridge, Mass: The MIT Press

Miller G A (1956) The Magical Number Seven, Plus of Minus Two: Some Limits on our Capacity for Processing Information. Psychological Review 63 (2): 81-96

Mirowski P (2002) Machine Dreams: Economics Becomes a Cyborg Science. Cambridge University Press, Cambridge, UK.

Muth J F (1961) Rational Expectations and the Theory of Price Movements. Econometrica 29 (3): 315

Myerson R (1999) Nash Equilibrium and the History of Economic Theory. Journal of Economic Literature 107 : 1067-82

Nelson R R (1994) Economic Growth via the Co-Evolution of Technology and Institutions. In: Leydesdorff L, Van den Besselaar P (eds) Evolutionary Economics and Chaos Theory: New Developments in Technology Studies Pinter, London, U.K.

Nelson R R, Winter S G (1982) An Evolutionary Theory of Economic Change. Belknap Press of Harvard University Press, Cambridge, MA.

Nielsen M, Chuang I (2000) Quantum Computation and Quantum Information. Cambridge University Press, Cambridge, UK.

Nyquist H (1924) Certain factors affecting telegraph speed. Bell System Technical Journal April 1924 : 324-346

Omnes R (1999) Quantum Philosophy: Understanding and Interpreting Contemporary Science. Princeton University Press, Princeton

Penrose R (1994) Shadows of the Mind. Oxford University Press, Oxford, UK.

Plott C, Sunder S (1978) Rational Expectations and the Aggregation of Diverse Information in the Laboratory Security Markets. Review of Economic Studies 45 : 133-53

Popper K R (1980) The Logic of Scientific Discovery. Routledge, London

Popper K R (1992) Realism and the Aim of Science. Routledge, New York, NY.

Prigogine I (1980) From Being to Becoming: Time and Complexity in the Physical Sciences. W.H. Freeman and Company, New York

Quine W V O (1969) Naturalized Epistemology. In: Quine W V O (ed) Ontological Relativity and Other Essays Columbia University Press, New York

Romer P M (1986) Increasing Returns and Long-Run Growth. The Journal of Political Economy 94 (5): 1002-1037

Romer P M (1990) Are Nonconvexities Important for Understanding Growth? American Economic Review 80 (2): 97

Rosen R (1991) Life Itself: A Comprehensive Inquiry Into the Nature, Origin and Fabrication of Life. Columbia University Press, New York, NY.

Rubinstein A (1998) Models of Bounded Rationality. The MIT Press, Cambridge, MA.

Scott M F (1989) A New View of Economic Growth. Oxford University Press, Oxford, UK.

Shannon C E (1948) The Mathematical Theory of Communication. Bell System Technical Journal 27 : 379-423, 623-656

Shannon C E, Weaver W (1963) The Mathematical Theory of Communication. University of Illinois Press, Urbana

Shapiro C, Varian H R (1999) Information Rules: a Strategic Guide to the Network Economy. Harvard Business School Press, Boston, MA.

Simon H A (1997) Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations. The Free Press, New York, NY.

Singh S (1999) The Code Book: The Science of Secrecy From Ancient Egypt to Quantum Cryptography. Anchor Books, New York, NY.

Smith J A (1991) The Idea Brokers. Free Press, New York

Stiglitz J E (1983) Information and Economic Analysis: a Perspective. The Economic Journal 95 : 21-41

Stiglitz J E (2000) The Contributions of the Economics of Information to Twentieth Century Economics. Quarterly Journal of Economics (140): 1441-78

Von Mises L (1969) Socialism: an Economic and Sociological Analysis. Jonathan Cape, London

Vromen J J (1995) Economic Evolution: An Enquiry into the Foundations of New Institutional Economics. Routledge, London

Vygotsky L S (1986) Thought and Language. The MIT Press, Cambridge, MA.

Weick K E (1995) Sensemaking in Organizations. Sage, Thousand Oaks

Weimer W, Palermo D (1982) Cognition and Symbolic Processes. Erlbaum, Hillsdale, NJ.

Williamson O E (1985) The Economic Institutions of Capitalism. The Free Press, New York, NY.

Zurek W H (1990) Complexity, Entropy and the Physics of Information. Perseus, Reading, MA.