A Rudimentary Theory of Information: Consequences for Information Science and Information Systems

PETROS A. M. GELEPITHIS

Kingston University, Kingston upon Thames, KT1 2EE, England.
E-mail: petros@kingston.ac.uk.

(Received October 30, 1996; accepted November 10, 1996)

The key fundamental notions of Information Science and Information Systems are 'information' and 'information system' respectively, with information the central common notion. Nevertheless, despite the huge literature on both 'information' and 'information system' their nature remains debatable and their conceptual nexus to each other and related notions like 'sign' and 'meaning' are, at best, fuzzy or incomplete.

The aim of this paper is to briefly review the notions of 'information' and 'information system' and to outline a theory of information. Specifically, in section one, we provide a brief analysis of 'information system' and a summary presentation of the major views on the nature of 'information', concluding that: (i) 'information' and 'communication' constitute the backbone of any theory of information; and (ii) all the relevant studies of information are fragmented, failing to provide a unifying theory and, in particular, to clarify the highly debatable nature of information. In section two, we provide the elements of a rudimentary theory of information and draw some of the consequences of this preliminary body of knowledge for Information Science and Information Systems.

KEYWORDS: information, information system, meaning, communication, understanding, theory

1 FOUNDATIONS OF INFORMATION SCIENCE AND INFORMATION SYSTEMS

The Information Science and Information Systems communities are known to stand quite apart from each other despite sharing some key foundational problems and despite the fact that the need

World Futures, 1997, Vol. 49, pp. 275–286
Reprints available directly from the publisher
Photocopying permitted by license only

for foundational and interdisciplinary work has been well established (see, for example, Machlup and Mansfield, 1983; Gitt, 1989; Checkland, 1992; Marijuan, 1996).

The key fundamental notions of Information Science and Information Systems are 'information' and 'information system' respectively, with information the central common notion. In addition, the considerable number of disciplines concerned with 'information' and 'information systems' has led to the development of a whole family of notions, closely related to that of information, (e.g., sign, symbol, meaning), which need to be clarified and become consistent with each other. This section provides a brief analysis of 'information system' and a summary presentation of the major views on the nature of 'information', concluding that: (i) 'information' and 'communication' constitute the backbone of any theory of information; and (ii) all the relevant studies of information are fragmented, failing both to provide a unifying framework, let alone a theory, and to clarify the highly debatable nature of information. We start with our analysis of the notion of 'information system'.

It is both well established and widely accepted that an information system is really a sociotechnical system.[1] Such a view makes clear the three *types* of fundamental notions required for its study. First, notions related to the concept of an 'information system' *itself*; second, notions related to all *those* (e.g., designers, managers, users) involved in the development of an 'information system' (we shall generically call those people *contributors*); and finally, notions related to the *tools* used in the development of an 'information system'. The following paragraphs present the particular sets of concepts characterising each of these three types and outline their links to the pair of *backbone* notions. We start with the system-related notions. These seem to fall in the following six, related, categories:

Group-1: Information–intelligence. The inclusion of information is, of course obvious; that of intelligence may be seen as less so to some people and hence a few words of explanation may be useful. A major category of information systems is those designed by humans. Its majority is due to its complexity and not its ubiquity in the universe. The complexity of an artificial information system, in turn, is due both to its links to the human elements of the designed system and to artificial systems processing information in ways which capture aspects[2] of human intelligent behaviour. Intelligence, therefore, in both its human and *emergent* machine form is necessary.

Group-2: Communication → Input-Output → interface (the arrow should be read as 'brings in the notion of'). It should be noted that 'communication' is necessary, above a certain threshold of complexity of the communicating entities. This should be juxtaposed with the *interacting* requirements of mere interfaces or input-output devices. Similarly, it is true that artificial intelligence systems continually approximate aspects of human systems and their number and penetration to new areas of human concern increases. The real challenge, then, is to interface and integrate artificial intelligence systems with human intelligence systems to develop complex human-machine systems. Communication is a must for the design, evolution, and effective and efficient running of such systems.

Group-3: Complexity → Hierarchy-? → emergent properties; and Group-4: Filtering → Hierarchy-? → emergent properties. Complexity is not very much[3] studied despite its characteristic importance for highly evolved natural systems and sufficiently richly-structured artificial or human-machine systems. Filtering is well advanced technologically but features pretty low in theoretical studies of both 'information' and 'information systems'. Both complexity and filtering bring in the notions of hierarchy and emergent properties, each of which raises fundamental issues of its own beyond the scope of this paper. The key link of both these two groups is with the notion of 'information system' rather than 'information'; more specifically, with the notion of a system's organisation. The reason for not including 'organisation' in the set of characterising notions is that it is a compound notion with components like complexity, and filtering.

Group-5: Goals, and control (including feedback). These two cybernetic notions remain centrally important for the study of information although not basic in the sense that 'information' and 'communication' are. As such they should play an important role in any *full* theory of information but they will not be included in our *rudimentary* theory.

Group-6: Design–Formalisability–Computability. This is an interesting group. Formalisability and computability are related, exclusively, to artificial information systems; design to both natural and artificial information systems. The former subgroup is closely related to the notion of uninterpreted system in formal studies but not *directly* related to information or the majority of its family notions, as they are defined in the next section. Design, in artificial information systems, is a process requiring communication (see next section for definitions and brief justification).

Contributors-related notions fall into two basic categories: (i) those involving the theoretical beliefs of a contributor; and (ii) those involving the non-theoretical beliefs of a contributor. The former category includes issues concerning the nature of organisation, society, science, and knowledge, as well as technical issues like computability, formalisability, and design. Essentially, we meet again here all of the notions characterising an information system itself, albeit mostly implicitly. Consider for example, 'knowledge' which requires a distinction to be drawn between individual and collective knowledge and hence brings in the issue of communication. Or, again, the nature of science which brings in the issues of formalisability and computability.

Finally, tools-related notions fall into three categories: (i) accuracy of representation; (ii); scope; and (iii) grain size. These are important, technical concepts which depend crucially on both the design and overall system requirements, and thus bring us via a third route to some of the basic notions introduced under the concept of system.

In summary, one can see that the two concepts which cut across all three types of fundamental notions required for the study of an 'information system' are information and communication and, therefore, these constitute the *backbone* of our rudimentary theory in the next section.

We come now to our summary presentation of the major views on 'information' and the few attempts made to provide a coherent framework for its related conceptual nexus. Concerning the nature of 'information' one may distinguish[4] seven major viewpoints. First, traditionally, information in terms of the probability of a signal (Shannon and Weaver, 1949).[5] Second, the conception of information as order (e.g., De Vree, 1996). Third, information in terms of knowledge and meaning at a mentalistic level (Langefors and Samuelson, 1976); and, more strongly, information as a mental not a material entity (e.g., Gitt, 1989). Fourth, information in terms of the notion of sign as a primitive (e.g., Stamper, 1985); Fifth, information conceived in terms, essentially, of the Popperian conception of the three worlds (e.g., Tully, 1985). Sixth, information in terms of truth conditions (see, e.g., Israel and Perry, 1990). Finally, information as a basic property of the Universe (e.g., Rzevski, 1985; Stonier, 1996); or, at least, as an objective commodity, or intrinsic to external objects (e.g., Dretske, 1981; Collier, 1990). Concerning the nexus of informational notions, the most notable attempt is that of the 'FRISCO group' who have set themselves the grand task of clearing the "conceptual foundations in the information system area", but so far[6] they have failed in developing a consistent framework that would be based on notions with a truly multidisciplinary acceptance. In summary, all the relevant studies are fragmented, failing to provide a unifying framework, let alone a theory, as well as a clarification of the highly debatable nature of information.

Taking together the above remarks on the notions of 'information' and 'information system', one is led to aim for a theory in the traditional sense of the word, that is, of a body of knowledge enabling an appropriate user to draw explanations and predictions about its subject matter as well as of controlling existing and designing new systems within its boundaries. This is a long list. The next section is confined: (i) to define 'information', 'communication', and the nexus of interrelated notions in a coherent and, if possible, unifying way which will minimise the vagueness of the notions involved as well as of the relations among themselves; and (ii) to draw some of the consequences of this preliminary body of knowledge for Information Science and Information Systems.

2 THEORY OUTLINE AND SOME CONSEQUENCES

We start with human[7] 'information', generalise to 'information', continue with the rest of the major family notions, and conclude with some clarificatory remarks with respect to our definition of meaning.

Human information $=_{df}$ Expressed human thought or set of human thoughts.

Human thought $=_{df}$ Set of human thought elements.

Human thought element $=_{df}$ Selected or prevailed neural formations.
Information $=_{df}$ Expressed thought or set of thoughts.
Thought of entity E $=_{df}$ Set of thought elements of E.
Thought element of entity E $=_{df}$ Selected or prevailed material formations of E.[8]
Symbol $=_{df}$ Human sign.[9]
Sign $=_{df}$ Configuration meaningful to a receiver.[10]
Signal $=_{df}$ Propagated configuration meaningful to a receiver.[11]
Data $=_{df}$ Potentially meaningful configurations.
The linguistic or perceptual meaning M of something s in the context C_s, for the entity E, at time t-symbol $M(s, C_s, E, t)$- is the *selected* or *understood* formations of the representational material of E, at t-symbol $R_e^{s,u}$.

To avoid potential misunderstandings, with respect to the last definition, the following three remarks are in order. First, the expressed meanings of an information system may be of the system itself or, equally well, those of another entity. For example, for a human perceived as an information system the meanings are internal to that human; for a present-day[12] computer though, the meanings it processes are those that some humans have chosen to represent in a computer processable form. Second, processing is very different from understanding. The former is akin to unconscious thinking and, in contrast to understanding, it may lead not to primitives (see below the definition of understanding and remarks on it). Finally, I have only presented here the generalised definition of meaning. For justification and discussion the reader is referred to Gelepithis (1989).

Now to the cluster of notions centred around communication and its basic constituent understanding. Although there is general agreement that 'communication' involves sharing and 'understanding' (see, for example, Cherry, 1957; Ogden and Richards, 1923; Rogers, 1986) no-one had really defined it until Gelepithis (1984). In what follows, we repeat those definitions, introduce the basic characteristics of the communication and understanding processes, and present a fundamental result that is used only to support consequences with respect to Information Systems.

Definition of communication: H_1 communicates with H_2 on a topic T if, and only if: (i) H_1 understands T {Symbol: $U(H_1 T)$}; (ii) H_2 understands T {Symbol: $U(H_2 T)$}; (iii) $U(H_1 T)$ is describable to and understood by H_2; and (iv) $U(H_2 T)$ is describable to and understood by H_1.

Definition of Understanding: An entity E has understood something, S, if and only if, E can describe S in terms of a set of primitives of its own.

The following characteristics of understanding and communication provide the basis of the consequences drawn next. First, understanding is structured. This has three aspects. One, being dependent on one's *own* primitives makes understanding dependent on time since *such* primitives do change with its passage. As an example, compare a toddler's primitives with those of a quantum physicist with respect to the notion of electricity.[13] Therefore, within one and the same person, understanding is 'layered' according to one's experience. Two, since understanding depends on *one's* own primitives, its end result, that is the individual knowledge reached, may well vary very significantly from person to person depending on the level of primitives reached by each person on a particular topic. Finally, understanding is structured as a consequence of the existence of two kinds of primitives: linguistic and sense primitives (Gelepithis, 1984; 1985). Second, understanding, if not immediate[14], requires a systematic approach to reach its objective. This follows directly from its defining characteristic of reducibility. Finally, understanding is not formal. This follows from the existence of the two types of primitives mentioned above.

On the basis of our definitions and characteristics of communication and understanding introduced above, a human, say H, and an intelligent machine, say M, would communicate on a topic T, expressible in language L; if and only if: either $P_H = P_M$ for T (P for primitive); or P_H and P_M could be described in terms of each other. Since linguistic primitives are reducible to sense primitives except if they are purely linguistic, one needs language to describe the senses and senses to understand language. Hence P_H and P_M could not be described in terms of each other. In other words, human-machine communication is impossible. This is a fundamental result, with ramifications extending beyond Information Science and Information Systems (for a full exposition of the argument and a general discussion see Gelepithis (1991). Here, it is used only to support consequences with respect to Information Systems.

The paragraphs of this section so far constitute our rudimentary theory of information. It is rather obvious that this preliminary body of knowledge is characterised by conceptual clarity, internal consistency, and a good degree of objective standing to a good number of the family notions related to 'information' and 'information system'. Next, we use such a rudimentary theory to derive some consequences for Information Science and Information Systems.

With respect to Information Science, the first consequence, derived from the nature of human information, is radical. Since human information is the expression of a set of selected or prevailed neural formations, there is no need for any new science of information; biology is perfectly adequate for the study of human or animal information. For information in general, physics takes up the role of biology.[15] Would the possible discovery of extraterrestrial information processors call for a science of information? I do not think so. The study of extraterrestrial information processors by humans, if possible, would only require the establishment of appropriate communication channels and the possible modification (including extension) of biological or physical principles. Naturally, the multidisciplinary and unifying perspectives which the proponents for a science of information advocate are laudable objectives which need to be adopted by biology or physics in their study of information. It is worthy of a note that the eventual, if not interrupted that is, emergence of machine intelligence will require the much closer cooperation of biology and Artificial Intelligence.

We turn now to some of the consequences with respect to Information Systems. First, complex human-machine systems *could not be fully formalised* except if all human elements were, eventually, to be replaced by artificial intelligence systems. The minimum number of human elements required to be kept in the system in order to be able to ascribe accountability to humans is a crucial, open question. Therefore, behaviour of such a system is in general non-computable. It can of course be constrained to produce only the computable aspects of its behaviour.

Second, since human-machine systems are non-computable, no general (i.e., system independent) information systems methodology can be constructed.

Finally, the specific methodologies for the development of knowledge-using human-machine systems are constrained by the processes of communication and understanding and therefore, cannot be purely formal. To design an effective and efficient information system it is necessary to include both formal and informal elements. This constraint and key methodological tool I call *the communication-understanding* principle. This last consequence can be seen clearly by considering the rationale of structured methodologies. It is based on two assumptions. First, that user requirements can be rigorously specified. Second, that such a specification will not include elements which are non-formalisable. But, user requirements are not even fully specified by the users. As a result: (i) the prerequisites for the use of structured methodologies do not hold true; and (ii) understanding of user requirements by the listener (be that a designer, a manager or whatever) differs from that intended to be communicated.

3 CONCLUSION

The work presented here is only a small part of what is required in developing a full theory of information and, equally important, presented only in outline or even in citation form, due to the usual paper-length restrictions. What is mostly required is to *consistently* put together as many of the various strands constituting the foundations of information as possible and to do that in a way that will be accepted by as many of the contributing disciplines as possible (eventually they should be *all* of course). The way forward is not for the faint-hearted.

Acknowledgements

Martin Robson has read through an earlier version of this paper and discussed with me some of the points made, or intended to be made. I hope I made good use of his revealing comments. Thanks

Martin. I would also like to thank the anonymous reviewer who made me rethink of my presentation on the basis of his/her astonishing and selectively useful comments.

Notes

1. For a well presented argument the reader is referred to Land (1985).
2. The question of autonomy and genuine intelligence of their *own* is beyond the purpose of this paper; the interested reader is referred to Gelepithis (1991).
3. Quite revealing in this respect is Simon's paper within the theoretical literature on the nature of complexity.
4. We exclude from our presentation all accounts which do not explicitly tackle the issue of the nature of information. For a full review of semantics covering the philosophical, linguistic, formal, and biological theories of meaning the reader is referred to Gelepithis (1988).
5. See also Kolmogorov (1968) for a common basis between probability and information theories.
6. I would like to note that the FRISCO group's work is under development and my criticism is based on their latest, but not final, public report (personal communication with IFIP WG 8.1 Task Group FRISCO, 1995).
7. In contrast to all other attempts, all my human-depended definitions are eventually cast in terms of neural (not necessarily neuronal) formations.
8. It may turn out that certain entities, exhibiting intelligent behaviour, may have 'thoughts' the nature of which is not captured by our definition. In such a case a decision will have to be made whether the scope of our definition needs to be modified, or it is preferable the discovered or designed entities to be classified as thoughtless entities with intelligent behaviour.
9. The most unified alternative view is Newell's (1990) based on the Physical Symbol System Hypothesis (Newell and Simon, 1976). For a summary review of the major views on the nature of symbols see Gelepithis (1995a).
10. Quite close to Charles W Morris' conception of sign (1939).
11. In sharp contrast to Shannon's theory (Shannon and Weaver, 1949).
12. For a discussion of intelligent machines see Gelepithis (1991).
13. For a discussion see Gelepithis (1995b).
14. That is, an intuition.
15. It should be noted that this does not imply a reductionist view. The issue of reductionism is much more complicated than it might appear from a face reading of the above sentence and although extremely interesting it is well beyond the scope of this paper.

References

Checkland, P. (1992). Information Systems and Systems Thinking: Time to Unite? In *Challenges and Strategies for Research in Systems Development*, W. W. Cotteman and J. A. Senn, eds, John Wiley and Sons.

Cherry, C. (1978). *On Human Communication: A review, a survey, and a criticism*. Third edition. The MIT Press.

Collier, J. D. (1990). Intrinsic Information. Hanson, P. P. (ed.). *Information, Language, and Cognition*. The University of British Columbia Press.

De Vree, J. K. (1996). A note on information, order, stability and adaptability. *BioSystems* Vol. 38, No. 2 and 3, pp. 221–227.

Dretske, F. I. (1981). *Knowledge and the Flow of Information*. Oxford, Blackwell.

Gelepithis, P. A. M. (1984). *On the Foundations of Artificial Intelligence and Human Cognition*. Ph.D. Thesis, Brunel University, England.

Gelepithis, P. A. M. (1985). *The Nature of Human Understanding: Human Primitives*. School of Information Systems Reports 1985–1988, Kingston University, England.

Gelepithis, P. A. M. (1988). Survey of Theories of Meaning. *Cognitive Systems*, No. 22, pp. 141–162.

Gelepithis, P. A. M. (1989). Knowledge, Truth, Time, and Topological spaces. *Proceedings of the 12th International Congress on Cybernetics*, pp. 247–256, Namur, Belgium.

Gelepithis, P. A. M. (1991). The possibility of Machine Intelligence and the impossibility of Human-Machine Communication. *Cybernetica*, Vol. XXXIV, No. 4, pp. 255–268.

Gelepithis, P. A. M. (1995a). *Artificial Intelligence: An Integrated, Interdisciplinary Approach*. (completed; with publishers).

Gelepithis, P. A. M. (1995b). Revising Newell's conception of representation. *Cognitive Systems*, Vol. 4, No. 2, pp. 131–139 (Special issue on Representation).

Gitt, W. (1989). Information: The Third Fundamental Quantity. *Siemens Review*, Vol. 56, No. 6.

IFIP WG 8.1 Task Group FRISCO (1995). *A Framework of Information System Concepts*. Personal communication.

Israel, D., and Perry, J. (1990). What is Information? Hanson, P. P., (ed.). *Information, Language, and Cognition*. The University of British Columbia Press.

Kolmogorov, A. N. (1968). Logical basis for information theory and probability theory. *IEEE Transactions on Information Theory*, Vol. 14, pp. 662–664.

Land, F. (1985). Is an Information Theory Enough? *The Computer Journal*, Vol. 28, No. 3, pp. 211–215.

Langefors, B., and Samuelson, K. (1976). *Information and Data Systems*. Petrocelli/Charte, New York.

Machlup, F., and Mansfield, U. eds (1983). *The study of information: Interdisciplinary messages*. John Wiley & Sons.

Marijuan, P. C. (1996). First conference on foundations of information science: From computers and quantum physics to cells, nervous systems, and societies. *BioSystems*, Vol. 38, No. 2 and 3, pp. 135–140.

Morris, C. W. (1939). Foundations of the Theory of Signs. *The International Encyclopedia of Unified Science*, Vol. 1, No. 2.

Newell, A. (1990). *Unified Theories of Cognition*. Harvard University Press.

Newell, A. and Simon, H. A. (1976). Computer Science as Empirical Inquiry: Symbols and Search. *Communications of the ACM*, Vol. 13, No. 3, pp. 113–126.

Ogden, C. K., and Richards, I. A. (1923*1956). *The Meaning of Meaning*. Harcourt Brace and Co. Inc. NJ, U.S.A.

Rogers, E. M. (1983*1986). Elements of Diffusion. (Extracts from Chapter 1 of *Diffusion of Innovations*, Rogers E. M. 3rd Ed, New York: Free Press, 1983).

Roy R., and Wield, D., (eds), *Product Design and Technological Innovation*, Open University Press.
Rzevski, G. (1985). On Criteria for Assessing an Information Theory. *The Computer Journal*, Vol. 28, No. 3, pp. 200–202.
Shannon, C. E., and Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press.
Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, Vol. 106, No. 6, pp. 467–482.
Stamper, R. K. (1985). Towards a Theory of Information. *The Computer Journal*, Vol. 28, No. 3, pp. 195–199.
Stonier, T. (1996). Information as a basic property of the universe. *BioSystems*, Vol. 38, No. 2 and 3, pp. 135–140.
Tully, C. J. (1985). Information, Human Activity and the Nature of Relevant Theories. *The Computer Journal*, Vol. 28, No. 3, pp. 206–210.