2009

Philosophy of Data (PoD) and its Importance to the Discipline of Information Systems

Brian Ballsun-Stanton
University of New South Wales

Deborah Bunker
University of New South Wales

Philosophy of Data (PoD) and Its Importance to the Discipline of Information Systems

ABSTRACT

In this document, we explore the Philosophy of Data (PoD) and its roots amongst other disciplines. The Philosophy of Data seeks to understand the nature of data through experimental philosophy. In order to understand the many different ontologies of data, information, and knowledge out there, this paper will describe part of the problem space in terms of other disciplines and make an argument for the establishment of this new philosophical field. Furthermore, we will show how the PoD is very important to IS scholars and practitioners.

Keywords

Philosophy of Data, Theoretical Basis of Information Systems, Data, Information, Knowledge, Ontologies, Multi-Disciplinary

INTRODUCTION

The many conflicting philosophies of data, information, and knowledge expressed and used in the discipline of Information Systems (IS) present a compelling problem to IS scholars. While each practitioner in this field has evolved his or her own philosophical understanding, based on necessity, evidence has demonstrated that there is little commonality between definitions across the entire discipline.

In order to investigate the nature data in ways that are both philosophically and structurally sound, we must not try to achieve a "one true" definition of data, for as different people use it, it means different things in each of their "realities". Furthermore, while technology is an expression of our field, the Philosophy of Data (PoD) must look at the nature of data, not its instantiation in devices.

This paper will seek to define part of the problem space of the PoD by exploring connections between areas that have a direct influence on it. Each of the related fields that work with and define data has concepts to share with the inherently multidisciplinary PoD, but this philosophy should stand on its own and the competing interests of other fields should not confuse the matter.

PoD – A MULTI-DISCIPLINARY PROBLEM-SPACE

Semiotics, the Philosophy of Information, Information Science, Information Technology (IT), the Philosophy of Technology, Education, Biology, the Philosophy of Science, and Information Theory all have their own unique weltanschauung to contribute to the foundations of PoD. Semiotics explores words and symbols as data. The Philosophy of Information explores the necessary sibling to data, Information. Information Science explores the formal manipulation of data as abstract symbol sets, introducing the vital component of Relations. IT continues the trend of relational modelling and explores functional data modelling and Information Theory explores the nature of the mechanical transmission of data. The philosophy of technology, meanwhile, contributes ethics and value-systems to data, as well as tools to help us probe its "fundamental philosophical nature." Education and Biology both explore the neurological basis of our brains' understanding of data. The Philosophy of Science's debate on facts also strongly informs the essence of data. These fields, while having different foci, all have concepts useful to contribute to the PoD. These fields relate to the multidisciplinary PoD within a number of domains shown in the diagram in Figure 1.

This paper will examine each disciplinary field to highlight the philosophical components that relate to the PoD. The Venn diagram in Figure 1 attempts to illustrate the rough topical areas of intersection and show the necessary multi-disciplinary nature of PoD.

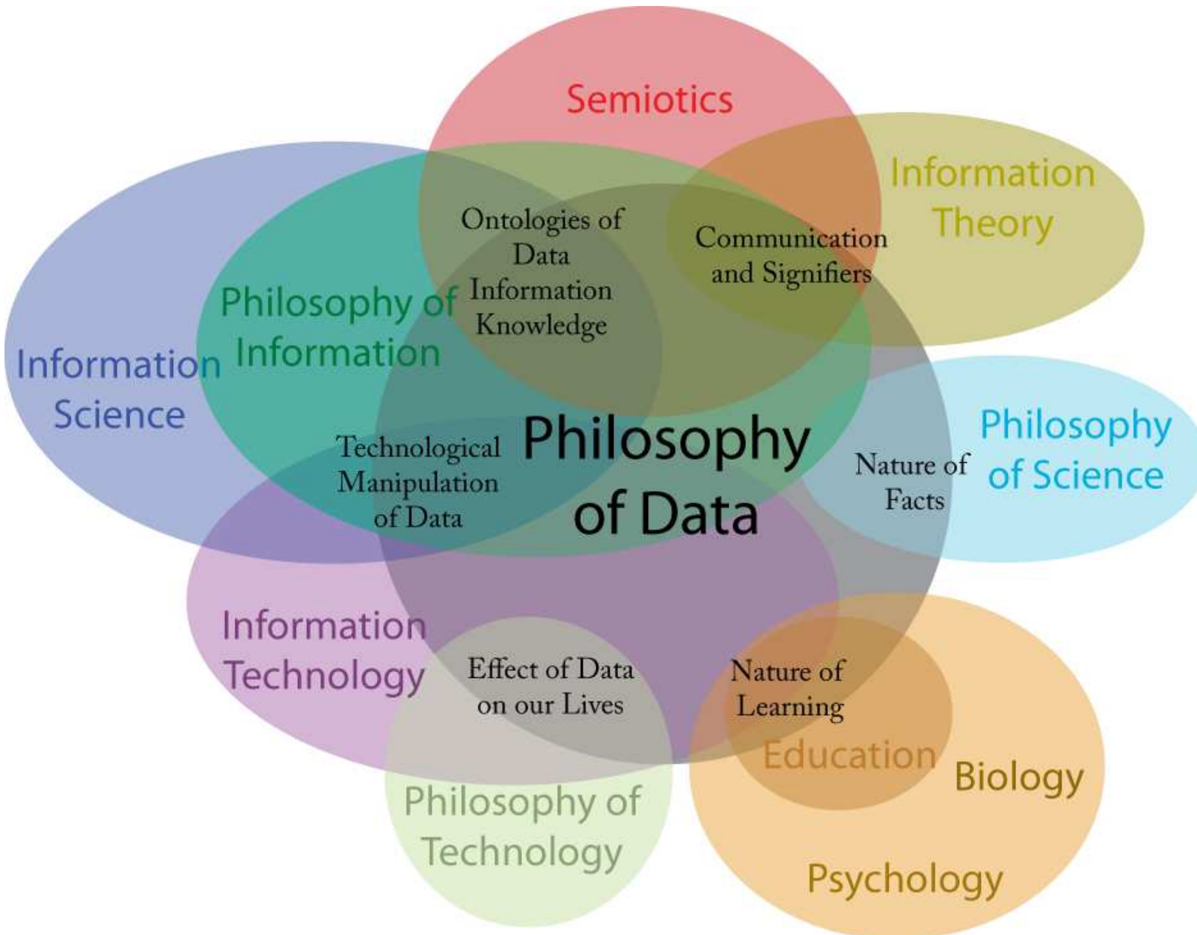

Figure 1. A Venn Diagram describing the associated fields and important topics of the Philosophy of Data.

Philosophy of Information

A closely associated field to the PoD is the Philosophy of Information. The Philosophy of Information, as defined by Floridi, is: "[T]he critical investigation of the conceptual nature and basic principles of information, including its dynamics (especially computation and flow), utilization and sciences; and the elaboration and application of information-theoretic and computational methodologies to philosophical problems." (Floridi 2002) This definition implies the existence of data, but also seeks to expand the Philosophy of Information's domain to Artificial Intelligence and into "solving" philosophical problems through computation. (Floridi 2004; Minsky 1986) While the PoD is mostly a subset of the Philosophy of Information, the goal of the PoD is to look at atomic units of information, not merely the computational uses of information. However, the Philosophy of Information does not provide us the contextual tools to explore interactions with data. By assuming that information is the first interesting aspect of the data-information-knowledge ontology, these philosophers may not see to the potential insights offered by investigating the interactions humans have with data.

Three authors of the Philosophy of Information discipline bear mentioning here. Dr. Floridi defines the Philosophy of Information as a discipline by positing a number of questions(Floridi 2004). These questions, useful to the Philosophy of Information qua itself, are also useful to this attempt to define the PoD. Here are some of his more pertinent questions: "What is Information?" In this simple question, he demarcates the Philosophy of Information just as the question: "What is Data?" defines the bounds for PoD. His second question is also highly applicable to the data/information dichotomy: "The I/O problem: What are the dynamics of information?" Here he explicitly disclaims investigation into information generation methodologies, choosing to investigate the information flow itself. The second question however, raises four fields of inquiry for the PoD. "How is data generated?", "How is data stored?", "How is data manipulated?", and "How is data perceived?" Looking once again at the central question of this paper, the desire to explore someone's interactions with data must answer all four questions.

Dr. Dodig-Crnkovic, on the other hand, calls for the increased practice of experimental philosophy when exploring these philosophical areas. As she claims the nature of philosophy is shifting, drawing it once more back into close proximity and operation with the other fields as, "It brings a potential for a new Renaissance, where Science and Humanities, Arts and

Engineering can reach a new synthesis, so very much needed in our intellectually split culture." (Dodig-Crnkovic 2003) Her paper can serve as a solid foundation for the arguments for using sociological inquiries to drive philosophy of real import and concern to everyday people, Information Systems practitioners, and PoD philosophers.

Brian Petheram discusses the tight connections between Information Systems modelling and the Philosophy of Information. He claims that, "by focusing on modelling as a key process of information systems development, … the deployment of something akin to a 'philosophy' is inevitable." (Petheram 1997) By exploring the Philosophy of Information with respect to data modelling, he also shows the tight binding between the philosophies of the designers and programmers of data models and the models themselves, thereby unconsciously influencing the end-users interactions with that data. This paper's key contribution is to demonstrate that there is an acknowledged link between philosophy and data, that this link is under-explored, and is worthy of exploration.

Information Science

Information Science, as a discipline, researches and defines the storage and processing mechanisms of what will eventually be information. In the field of Information Science, the PoD is concerned with information processing/retrieval (Relational Algebra), the creation of data-information-knowledge ontologies, relational algebra, and data modelling. Relational Algebra, popularized by E.F. Codd in 1969 (Darwen and Date 1995) is the foundation of modern relational databases. Relational Algebra defines a mathematical model for the manipulation of discrete sets to get the answer required through manipulating relationships.

By modelling reality into relationships suitable for a database, designers and programmers must perform certain abstractions and generalizations. Without considering the future implications of these abstractions in the philosophy of the data model, it is possible to cause a disconnection between reality and the reality represented by the model. Codd's seminal paper arguing for the superiority of the relational model over the hierarchical and network models of database design has two specific contributions (Codd 1969): it defines a relation and relationship in a mathematical sense, which demonstrates how data can be contextualized with other data by providing application of set operations to sets defines as relations and relationships. However, the more subtle, philosophical contribution is in its normative recommendations for a "normal form". Codd claims: "There are usually many alternative ways in which a relational model may be established for a data bank. In order to discuss a preferred way (or normal form), we must first introduce a few additional concepts … and establish some links with terminology currently in use in information systems programming." (Codd 1969) This claim defines the philosophical field of data modelling by making assertions about correct and incorrect ways of modelling data, but fails to consider the philosophical implications of so doing.

C. J. Date, another important author, explores data modelling, the description of the fundamental sets used in Relational Algebra (Date 1986). Data modelling contains two very important and distinct sub-fields: data requirements gathering and requirements modelling. The practice of requirements gathering tries to understand the universe of discourse (a delineated portion of reality relevant for modelling) under discussion. The requirements modelling process then applies that understanding of reality to a synthetic data model wherein data representing reality can be stored.

The Doctors Marcos started another useful discussion in the field of data modelling; they explore the question of the double meaning of the term model in the computing fields:

> "The design of the Database is crucial to the process of designing almost any Information System and involves two clearly identifiable key concepts: schema and Data model, … the term "model" is commonly applied indistinctly to both, the confusion arising from the fact that … the notion of "model" has a double meaning of which we are not always aware. If we take our idea of "model" directly from empirical sciences, then the schema of a Database would actually be a model, whereas the Data model would be a set of tools allowing us to define such a schema." (Marcos and Marcos 2001)

This assertion has a direct bearing on the PoD as the Marcos not only seek to define the process as data modelling, and the result of the process as a schema, but are also exploring the philosophical implications of both the tool of the schema and the technique of the data model.

Interestingly, Information Science and the Philosophy of Information intersect when they attempt to delineate the differences between data, information, and knowledge by creating an ontology. Dr. Zins (Zins 2007) recently conducted research on behalf of the Israel Science Foundation attempting to map all the currently extant ontologies. His collection of one hundred and thirty definitions explores three different forms of ontology: "Interrelations", "Information versus Knowledge", and "Synonyms". Zins' survey of the field has a number of important repercussions for the PoD, especially exploring how the PoD relates to the other philosophies. We can consider this research to be one of the core components of the PoD as it is gathering experimental evidence to differentiate data from information and knowledge. However, this research does not consider the philosophical implications of the positions it has surveyed.

Specific mention is required of Dr. Tuomi's reverse hierarchy view. She observes that the classic ontology of facts as data which become information which becomes knowledge is rather the opposite, specifically that: "Data emerge after we have information and that information emerges only after we have knowledge." (Tuomi 1999) The reverse hierarchy is notable for a number of reasons. Primarily, it explores a causal view of data instead of a component view. The component view, described above in the traditional sources is that data is a subset of information (or information is data with extra stuff bolted on), and that information is a subset of knowledge. The traditional view does not ascribe requirements to any of the three outside their simple existence. The causal view, as express by Tuomi, describes knowledge as that which can cause the creation of information, and likewise information to data. Not only does knowledge create context for information or a change in granularity, but it also creates the need for that information. Thus, Tuomi indirectly reminds us that temporality and causation are important in any discussion of PoD, due to the difference in definition perspective.

Semiotics

While the Philosophy of Information is one of the parents, the other field that could be termed a parent of the PoD is that of Semiotics (de Saussure 1985). Sassure defines Semiotics, the study of signs, as:

> "… [T]he role of signs as part of social life. It would form part of social psychology, and hence of general psychology. ... It would investigate the nature of signs and the laws governing them. …. Linguistics is only one branch of this general science." (de Saussure 1986)

As the form and meaning of data strongly relates to its nature, a large component of the PoD must explore the semiotic aspects of data. With that said, however, Semiotics is not the entirety of the PoD as the focus of semiotics is on the nature of communication and the interpretation of signs for meaning. Despite the generation of meaning from signs being central to the PoD, limiting our investigations to human sign-creators ignores the universe of artificially created data that is just as or more meaningful than the signs humans create for themselves.

Eco, in his research in Semiotics, explores a number of items of significance to the PoD. First among them is the coding of signs, the way we represent meaning in symbols. Eco first describes this coding process when he explores Pierce's definition of semiotics. He states: "A sign can stand for something else to somebody only because this 'standing-for' relation is mediated by an interpretant." (Eco 1979) As data are signs (but not all signs are data, perhaps) this aspect of Eco's understanding of semiotics represents another crucial element the intended study. To restate the aim, "How does someone's PoD influence their interactions with data?" part of the field should explore how someone misunderstands data produced by someone or something else. The way of understanding people misunderstanding data is through Eco's discussion of interpretants.

The other significant exploration in semiotics that is relevant to the PoD is Galison's exploration of "The Context of Context". He explains the nature of trading zones to enable communication between different types of scientists in the field of physics:

> "... the subcultures of Physics are diverse and differently situated in the broader culture in which they are prosecuted. But if the reductionist picture … fails by ignoring this diversity, a picture of physics as a merely an assembly of isolated subcultures also falters by missing the felt interconnectedness of physics as a discipline. ... I repeatedly use the notion of a trading zone, an intermediate domain in which procedures could be coordinated locally even where broader meanings clashed."(Galison 1997)

The concept of the trading zone should not only be applied to Physics (Baird and Cohen 1999; Derry, Gernsbacher and Schunn 2005). On the contrary, when we apply the trading zone to the PoD, we can see that as people and machines must construct implicit and explicit trading zones to communicate data, opportunities for miscommunication arise. Therefore, the trading zone should be one of the primary things to explore when exploring organisational and international communication of data as well as the encoded trading zones of machine-stored data. Furthermore, the conscious awareness of a trading zone is also a very important investigation into a person's specific understanding of the nature of Data.

Philosophy of Science

The next most significant contributor is the Philosophy of Science. While the PoD has no interest in how science is conducted, it is a clear recipient of the repercussions of the "What is a fact?" The debate carried out by Popper (Popper 1959), Kuhn (Kuhn 1970), Lakatos (Lakatos 1970), and Feyerabend (Feyerabend 1993) on the nature of fact is central to science and as facts are generally represented as data or agglutinations of data. As scientists and philosophers have not "solved" the philosophical problem presented by the question of fact, the debate must therefore carry over to the PoD. Facts as data are represented by Carnap (Carnap 1946) and Popper as observations. The dispute then changes into what or who can create facts/data. The nature of context of the facts supporting theories then evolves past Kuhn and Lakatos' paradigms or research programmes into requiring that some data is meaningless when placed with others. The anarchistic approach of Feyerabend upsets the equilibrium even further by disputing the very nature of science. When considering the requirements of validity it is important to observe how people equate data with factual claims.

Philosophy of Technology

While the Philosophy of Science has much to say about the nature of fact, the Philosophy of Technology explores how technology mediates our daily lives. Of significance is the mediating nature of technology on data. Whether technology is the observer, recorder, manipulator, or other actor upon data, it is centrally bound to the nature of data. We must consider at the very least three perspectives on the mediating influence of technology: John Dewey (Dewey 1997) and Martin Heidegger (Heidegger 1977) on the future of technology, Dreyfus (Dreyfus, Dreyfus and Athanasiou 1986) and Borgmann on the role of technology in life and Artificial Intelligence, and Marshall McLuhan's work on the nature of the technological medium.

Blattner's exploration of Dewey and Heidegger's positions of understanding, cognition, and technology is very striking and relates to the very heart of PoD. He explores how, "understanding is primarily a sort of practice, or as Heidegger says, a sort of competence or ability, and theoretical knowing is a part of this." (Blattner 2000) As technology is both realized data and the vehicle for the realization of data, Blattner's exploration of Dewey and Heidegger is very important because it allows for inferences of how data interacts with understanding and the understanding engineered into technology as embodied practice.

Borgmann's discussion between "Efficiency for Efficiency's Sake" and "Focal Points" shows the importance of data and information to philosophers of technology. On the nature of data and reality, he asserts:

> "Yet reality is not divisible into structure and contingency without remainder. That the world is woven together is a further contingency of lawlike necessity, and so on endlessly even through intelligibility. Reality is both knowable and unsurpassable. Positing Platonic structures is the attempt to get control of reality by dividing it all the way down..." (Borgmann 1984)

We must investigate if these platonic structures are data. Thus, to the philosophers of technology, questions of how data interacts with reality through the medium of technology are vital to the field.

We are also interested how the Philosophy of Technology explores the nature of communication, specifically in McLuhan's exploration of the nature of the medium, he notes:

> "In a culture like ours, long accustomed to splitting and dividing all things as a means of control, it is sometimes a shock to be reminded that, in operational and practical fact, the medium is the message. This is merely to say that the personal and social consequences of any medium -- that is any extension of ourselves -- result from the new scale introduced in our affairs by each extension of ourselves." (McLuhan and Fiore 1967)

The concept of extension of self via media, being how each communication through a medium itself carries meaning just as the message contained in that communication does is a very interesting concept to apply to the PoD. Specifically, inasmuch as various media encode data differently, we cannot simply consider data as a platonic unit. Both Semiotics and the Philosophy of Technology would have us be mindful of the fact that the encoding of the data, much less the context or domain, carries meaning itself. Thus, when trying to explore people's understanding of the nature of data, we must be mindful of what media they ascribe to data storage, and what perceptions they have of the medium's effect on data.

Information Technology

On the theoretical side of things, we open with Information Theory where Shannon founds a Mathematical Theory of Communication. He opens this field with a dedication to the engineering as opposed to semantic side of communication: "The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or correlate with some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem." (Shannon 1951) It is vital that the PoD considers the nature of data transport and those associated fields on data itself. Moreover, by introducing the concepts of entropy, quantifiable error, and the correction thereof, Shannon and his followers add another dimension to the Philosophy of Information: physicality of the intangible.

Science

The natural sciences may also have a large contribution to the PoD. Specifically, we find it in the intersection of the fields of Psychology, Biology, and Education. The intersection of Psychology and Education gives us an exploration of the intentional imparting of knowledge (whatever we actually end up understanding that to be), supplies theories like Piaget's theories of learning (Piaget, Piattelli-Palmarini and Chomsky 1980), alongside Chomsky's theories of language(Chomsky 1992). The study of the neurological (nature) and psychological (nurture) basis of the way we interact with data must strongly inform any philosophy thereof.

CONTRIBUTIONS TO INFORMATION SYSTEMS

With all these contributing fields, the discussion now centres on the question how PoD contributes to the discipline of Information Systems. What components of PoD are useful for IS study, and why? There are three major contributions from PoD here. An awareness of the relative, transitive, nature of our own, personal, definitions of data, a deepened understanding of the contextual nature of information systems implementations, and an awareness of the impact of cultural differences in the construction of data.

The awareness of the limitations of our own definitions is paradoxical: how can we perform scholastic research if we cannot be certain of the fundamental terms of our field? The PoD helps here by containing the problem space wherein most potential definitions of the term lie. By pinpointing the exact definition of the term under examination within the problem space, not only do we create a level of commensurate comparison between different types of data, but we also allow other people to more accurately understand the term in the way the researcher is using it, and thereby minimize miscommunication.

Using PoD to realize the pitfalls of the contextual nature of IS implementations is another key contribution. While some organisations may have the luxury of hiring designers to create a completely custom-built system for their particular problem, even that system will have to interact with systems built by other people in different contexts. While published standards define these systems to some extent, similar functionality, and suchlike, there is a limited appreciation for the different definitions of data: most, if they consider it at all will merely try to design for the little-endian or big-endian architecture of their target system. However, with Electronic Support Systems, Databases and objects, and the global nature of Electronic Commerce and the Internet, the way these systems conceptualize, process, store, and present data for understanding will be remarkably different. While the major breaks in understanding the nature of data are easy to detect as language barriers, translation, and internationalization can create major flaws in database design and UI presentation, more subtle flaws between intent, design, and use can creep in. Errors due to these flaws tend to be attributed to other problems, "poor design" or "poor training" for the most part. Looking at a study exploring the success of LINUX design (Augustin, Bressler and Smith 2002) one problem to be overcome was to create a synthetic culture of sharing. While the artificial creation of context through a new culture works in terms of the creation of an operating system, not all user bases should adapt themselves to their to the encoded databases' perspective of reality. Thus, by exploring these contexts within the framework developed by PoD philosophers, solutions that bridge contexts may suggest themselves.

The third component informs the above stated problems of context: different cultures at whatever granularity: organizational, national, educational. These cultural differences, apart from the problems created by the context of the creation of the IS product, can pose major obstacles. Cultural differences of the understanding of data may even express themselves across the culture of an organization. If we consider Intelligence Agencies in the United States government, it is easy to imagine the culture of the intelligence gatherers, the "data creators" being vastly different from those who interpret data and even more distant from the politicians who consume data (Gookins 2008). Using Schein's model of Organizational Culture (Schein 1984), we can form a clearer view of this process. Assumptions inform the values of each node of the culture, and the artefacts reflect the culture's interpretations of the values upon the tacit assumptions. The difficulty is that we then read those assumptions into the artefacts that the culture consumes, even if the origin of those artefacts was from a different culture. Thus, using the PoD to explore how culture informs data will help groups/cultures (organizations) reduce communication errors introduced through misunderstanding a different culture's philosophy of data.

CONCLUSION

These domains, authors, and fields that explore ideas and philosophy behind data and its definition and use, are all extremely useful and valid in their own right. Additional understanding of the PoD occurs, however, when scholars and philosophers explore these domains in the development of a multidisciplinary PoD. While this paper must be brief, the partial articulations of each field present, in combination, a convincing argument for the need to explore the PoD as a whole and some of the theoretical boundaries that differentiate this discipline from its roots.

ACKNOWLEDGEMENTS

We would like to thank Dr. Fouad Nagm for his assistance in early conceptual editing and suggestions for the Venn diagram.

REFERENCES

1. Augustin, L., Bressler, D. and Smith, G. (2002) Accelerating software development through collaboration, ACM New York, NY, USA, 559-563.
2. Baird, D. and Cohen, M. S. (1999) Why trade?, *Perspectives on Science*, 7,2, 231-254.

3. Blattner, W. D. (2000) The primacy of practice and assertoric truth: Dewey and heidegger, in M.A. Wrathall, J. Malpas and H.L. Dreyfus (Eds.) *Heidegger, authenticity, and modernity: Essays in honor of hubert dreyfus*, MIT Press, Boston, 231-250.
4. Borgmann, A. (1984) Technology and the character of contemporary life: A philosophical inquiry, University of Chicago Press.
5. Carnap, R. (1946) Theory and prediction in science, *Science*, 104,2710, 520-521.
6. Chomsky, N. (1992) On cognitive structures and their development: A reply to piaget, in B. Beakley and P. Ludlow (Eds.) *The philosophy of mind: Classical problems/contemporary issues*, MIT Press, Boston, 393-396.
7. Codd, E. F. (1969) A relational model of data for large shared data banks, *Communications of the ACM*, 13,6, 377-387.
8. Darwen, H. and Date, C. J. (1995) The third manifesto, *SIGMOD Rec.*, 24,1, 39-49.
9. Date, C. J. (1986) Relational database: Selected writings, Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA.
10. de Saussure, F. (1985) The linguistic sign in semiotics: An introductory anthology,(ed) r. Innis, *Indiana University Press, Bloomington, IN*, 24, 46-46.
11. de Saussure, F. (1986) Course in general linguistics, Open Court.
12. Derry, S. J., Gernsbacher, M. A. and Schunn, C. D. (2005) Interdisciplinary collaboration: An emerging cognitive science, Lawrence Erlbaum Assoc Inc.
13. Dewey, J. (1997) How we think, Dover Publications.
14. Dodig-Crnkovic, G. (2003) Shifting the paradigm of philosophy of science: Philosophy of information and a new renaissance, *Minds and Machines*, 13,4, 521-536.
15. Dreyfus, H. L., Dreyfus, S. E. and Athanasiou, T. (1986) Mind over machine: The power of human intuition and expertise in the era of the computer, The Free Press.
16. Eco, U. (1979) A theory of semiotics, Indiana University Press.
17. Feyerabend, P. K. (1993) Against method, Verso.
18. Floridi, L. (2002) What is the philosophy of information?, *Metaphilosophy*, 33,1/2.
19. Floridi, L. (2004) Open problems in the philosophy of information, *Metaphilosophy*, 35,4, 554-582.
20. Galison, P. L. (1997) Image and logic: A material culture of microphysics, University Of Chicago Press.
21. Gookins, A. J. (2008) The role of intelligence in policy making *SAIS Review*, 28,1, 65-73.
22. Heidegger, M. (1977) The question concerning technology, *The Question Concerning Technology and Other Essays*, 3-35.
23. Kuhn, T. S. (1970) The structure of scientific revolutions, University of Chicago Press Chicago.
24. Lakatos, I. (1970) Falsification and the methodology of scientific research programmes, *Criticism and the Growth of Knowledge*, 4, 91–195.
25. Marcos, E. and Marcos, A. (2001) A philosophical approach to the concept of data model: Is a data model, in fact, a model?, *Information Systems Frontiers*, 3,2, 267-274.
26. McLuhan, M. and Fiore, Q. (1967) The medium is the message, Bantam, New York.
27. Minsky, M. (1986) The society of mind, Simon & Schuster, Inc. New York, NY, USA.
28. Petheram, B. (1997) Backing into philosophy via information systems, in R. Winder, S.K. Probert and I.A. Beeson (Eds.) *Philosophical aspects of information systems*, Taylor & Francis, Inc., 113-116.
29. Piaget, J., Piattelli-Palmarini, M. and Chomsky, N. (1980) Language and learning: The debate between jean piaget and noam chomsky, Routledge, London.
30. Popper, K. R. (1959) The logic of scientific discovery, (First English ed.), Hutchinson & Co.
31. Schein, E. H. (1984) Coming to a new awareness of organizational culture, *Sloan Management Review*, 25,2, 3-16.
32. Shannon, C. E. (1951) Prediction and entropy of printed english, *Bell System Technical Journal*, 30,1, 50-64.
33. Tuomi, I. (1999) Data is more than knowledge: Implications of the reversed knowledge hierarchy for knowledge management and organizational memory, *J. Manage. Inf. Syst.*, 16,3, 103-117.
34. Zins, C. (2007) Conceptual approaches for defining data, information, and knowledge, *Journal of the American Society for Information Science and Technology*, 58,4, 479-493.