A Semiotic Perspective on Semantic Interoperability

Karin Ahlin
PhD Student Informatics
Mid Sweden University
karin.ahlin@miun.se

Ted Saarikko
Assistant researcher
University of Gothenburg
ted.saarikko@ituniv.se

Abstract
While information technology enables us to access more material than ever before, we need to come to grips with the disagreeable fact that data is not information. Sharing data without safeguarding comprehension may lead to confusion at best and disaster at worst. The issue at hand is one of ensuring semantic interoperability between actors from disparate contexts. Research into this problem is plentiful, but typically focused around specific subject matters, limiting its appeal to a limited range of scholars and practitioners. Based upon a literary study, we identify two extreme approaches to managing semantic interoperability. These are denoted top-down and bottom-up. We illustrate real world instances of these approaches using the TFI-model based on a case study encompassing two organizations. Our theoretical model is found to be a viable lens through which to generalize and interpret issues pertaining to semantic interoperability between human actors. We therefore see a need for further research into human-based on semantic interoperability.

Keywords: Semantic interoperability, top-down, bottom-up, TFI-model, semiotics

1 Introduction
As our ability to share and process data increases, we fancy ourselves on the cusp of realizing the vision of the ubiquitous information society. However, we must face the disagreeable fact that data is not information and sharing does not guarantee understanding. We must learn to distinguish between quantity (data) and quality (information) if we are ever to resolve this dilemma. Overconfidence in our ability to transfer data can lead to serious consequences – perhaps the most spectacular of which being the Mars Climate Orbiter that suffered catastrophic failure due to application of English units rather than metric units in one of its data files (NASA, 1999).

The issue of conveying information rather than data has already received significant attention under the guise of semantic interoperability. Unfortunately, research into semantic interoperability tends to suffer from one of two conditions. It is either rather technically biased (Backhouse & Halperin, 2009), or it is based in a specific context which is imbued with a nomenclature that is often quite esoteric. A search for “semantic interoperability” in research database such as Science Citation Index Expanded yield a large number of hits pertaining to eGovernment, eHealth, semantic web and Geographic Information Systems. Each of these areas approaches the difficulties of semantics from their own perspective and using their own ontology. The current trend of imbuing semantic interoperability with a technical language puts it closer to technical knowledge than scientific knowledge, thus
limiting its potential for verification or falsification to a limited community (Boisot, 1995, p.72-73; Popper, 1974, p.81). From a management perspective, the lack of accessibility means that more precious time has to be spent absorbing data (Rockart, 1979), not to mention the risk of finding oneself at the mercy of technical experts (Ackoff, 1967; Ross & Weill, 2002). Hence, we see that there are advantages to approaching semantic interoperability from a general perspective to the scholar as well as the practitioner.

Our intent is to approach semantic interoperability from an informatics perspective by studying how it may be discussed from a semiotic perspective. Therefore, we treat semantic interoperability not as a state of technical compatibility, but rather the ability of individuals to derive the same meaning from a set of data. Based on our starting point, literature suggests two diametrically opposed views on this topic; we refer to these as top-down and bottom-up. In addition to semiotics, we describe these approaches utilizing the TFI-model which distinguishes between technical, formal and informal aspects of the organization. An empirical study is then undertaken in an effort to demonstrate what form our theoretical posits may assume in practice. This is not intended to validate the premise presented, but rather to provide an illustration much in the same manner as Hirschheim & Klein (1989).

2 Semiotics

Advances into information technology have taken us from a situation where data was effectively tied to a limited geographical area, to one where data may be transmitted in vast quantities to anyone or everyone (Orman, 1983). However, the increase in distance (physical as well as cognitive) between transmitter and receiver has highlighted the inherent difficulties in correctly interpreting data out of context (Langefors, 1973, p.242-249; Liebenau & Backhouse, 1990, p. 25-27; Boisot, 1995, p. 93-115; Harvey, Kuhn, Pundt, Bishr & Riedemann, 1999).

While computers and digitized data exacerbated this issue through sheer volume, it is by no means a novel topic. The same basic issue has been extensively studied under the guise of semiotics for the better part of a century (Liebenau & Backhouse, 1990, p. 15). Semiotics, of which semantics are a subset, is the study of signs and how they facilitate communication. Semiotics may be divided into four components: Empirics, syntactics, semantics and pragmatics (Liebenau & Backhouse, 1990, p.11-79). Empirics and pragmatics may be further subdivided (Stamper, Liu, Hafkamp & Ades, 2000) if needed, but that level of detail may be considered redundant for the discussion at hand. Empirics form the basic physical components of transferring data – the medium that facilitates transmission from sender to receiver. Syntactics provides us with rules that enable us to impose some manner of structure upon the data which we transmit or receive. It is only with an appropriate structure that we may process and refine data regardless of whether we are compiling financial data or digitizing a painting. Semantics pertain to the transfer of intended meaning. The meaning of data is often tightly linked with a particular context (Magoulas & Pessi, 1998, p. 366-369; Harvey et al, 1999). Thus, the focus of semantics is the relationship between what is being transmitted and what is being understood. One of the basic tools with which to accomplish this is a common terminology (Liebenau & Backhouse, 1990; Holsapple & Joshi, 2002).
Pragmatics represent the manner in which understanding prompts action. Even with mutual understanding assured, we cannot unconditionally assume that a piece of information will prompt similar action in two different individuals (Star & Bowker, 2002).

Overcoming the semantic barrier, such as by common ontology, enables organizations to bring disparate skill-sets to bear on complicated tasks such as development projects (Boland & Tenkasi, 1995). Ontologies may be created in several fashions, ranging from being the brain-child of one person to extensive collaborative efforts (Holsapple & Joshi, 2002). However, establishing a shared ontology is a delicate and lengthy process which often runs counter to the rational desire to economize on information processing (Boisot, 1995, p. 39-82). Indeed, existing organizational channels of communication may prove detrimental to establishing a new ontology or new ways of perceiving one’s surroundings (Boland & Tenkasi, 1995; Miller, 1993).

3 The TFI-model

In order to frame our discourse, we find it imperative to draw upon a plethora of sources in order to do the subject justice. In addition to semiotics, we intend to utilize the TFI-model (Liebenau & Backhouse, 1990, p. 109-112; Stamper et al., 2000) which in broad terms outlines the interplay between organizational layers and systems without constricting our discussion.

![Figure 1: TFI-model adapted from Stamper et al. (2000)](image)

The constituent parts of the TFI-model are technical systems, formal organization and informal organization. Technical systems are, in a word, artefacts. While hardware and software might be the most common artefact in conjunction with information systems, they are not the only inhabitants of this domain. In essence, anything that contains or conveys information may be considered an artefact. Formal organizations are specified patterns of action – usually by design. While these patterns may be distributed by means of an artefact (such as a manual), essence of formal systems are the codified and preordained nature of the actions that are carried out. Lastly, informal organizations are structures that are more happenstance than designed. That being said, they may of course to some degree be anticipated based on our individual qualities such as skill-set, experience and disposition.
Furthermore, an informal status does not preclude these actions from being considered common or natural in a particular context.

The TFI-model posits that these three domains are layered in a manner that technical systems are subsumed under formal organizations – which in turn are subsumed under informal organizations. Hence, a technical system requires one (or more) formal organizations (such as rules or standards) in order to serve any functional purpose. In much the same manner, a formal organization has to be accepted by the informal organization (such as norms and values) in order to gain any footing.

4 Managing semantic interoperability

Based on our literature study we have identified two extreme approaches to semantic interoperability. We refer to these as top-down and bottom-up.

4.1 Top-down

The expression top-down in itself is widely used in conjunction with change management where senior management drives changes based primarily on a strategic view of the organization (Kirkbride, 1993; Nadler & Tushman, 1997, p.52-54). The decision to adopt the term top-down in our discussion on semantic interoperability is based on the locus of this logic – not the mechanics itself. One must make the important distinction between change management and semantic interoperability in that the former deals with action, the latter comprehension. It is of course quite possible that there may be a temporal or even causal relationship between the two areas – such as insufficient comprehension prompting change efforts – but such a discussion is beyond the scope of this paper. Furthermore, we must distinguish between the ownership of the underlying logic for structuring data and the data itself. A unit or subdivision may control data regarding its operations, customers and personnel in that local staff has the right to add, edit or delete what is stored in local/regional information systems. This does not mean that the unit or division is allowed to alter the way in which data is structured (Ross, Weill & Robertson, 2006, p. 28-38). Authority to amend the data structure may reside far removed from the unit(s) in which it is implemented.

A top-down approach to semantic interoperability expresses an implementation of formal standards that are intended to promote communal action throughout the organization. Standards are disseminated or even enforced by means of artefacts of some kind – such computerized information systems. The standards in question range from management policies to regulatory imperatives (DiMaggio & Powell, 1983).

The logic underlying the top-down approach is to a large extent rooted in Simon’s (1962) notion of the organization as a goal-seeking entity. The organization seeks to achieve a highly complicated objective that must be broken down into manageable steps in order to be managed by an organizational unit. Thus, the manager – being the designer of the organization – is concerned with not just what is to be done, but also how things are to be done in order to optimize the organization as a whole (Simon, 1996, p. 4-5, 110-120; Simon, 1997, p. 186-197). In semiotic parlance, the designer is primarily concerned with pragmatics as he/she alone possesses an understanding of the system as a whole. The most common expression of this is how expertise and know-how is assigned to tasks; a common expression of this being hierarchies such as a bureaucracy or a multi-departmental form (Tompkins, 1987; Chandler, 1962, p. 325).
Davenport & Prusak (1997, p. 74-75) expresses the top-down perspective as a “monarchy” where one person or unit in the organization determines the shape and form of information processing. There are certain benefits associated with centralization, such as efficiency (Boland & Tenkasi, 1995) and standardized terminology (Holsapple & Joshi, 2002). However, the limited world view of a handful of individuals may also bring about detrimental effects, including poor information quality (Ackoff, 1967), arbitrary decision making (Ciborra, 2000, p. 39) and insufficient understanding of ends and means (Peppard, 2007).

4.2 Bottom-up

The bottom-up perspective may be thought of as a task-oriented approach to governance where the term “management” implies coordination rather than control. While there are several nuances in how people perceive the concept of bottom-up (Sabatier, 1986; Kirkbride, 1993), we intend to look upon this as a state where each organizational unit has their own taxonomy which is suited for their needs. Davenport & Prusak (1997, p.72-74) approximates this concept in their decentralized notion of “feudalism” where cooperation between departments is atypical. Thus while information structures corresponds to each organizational unit, top-level management may perceive information redundancy as high and have trouble accessing specific pieces of information without the aid of localized staff.

Organizations strive towards certainty in their decision-making despite the insurmountable complexities of the real world (Galbraith, 1973, p. 4-6). Failing to complete this Sisyphean task, organizations adapt by isolating their core technologies from uncertainty or ambivalence (Thompson, 1967, p.10-13). In doing so, one may formulate an operational logic that focuses on the task itself and perceives the surroundings only in the simplest of terms. It is then up to various supporting functions – such as management – to ensure that the core processes may operate under this premise.

The logic underpinning this perspective is described by Churchman (1971, p. 53-68) as the principle of non-separability. According to this view, the constituent parts of a system may not be designed or analyzed in isolation. Doing so would omit relativistic properties that may only be ascertained in a given context. Consequently, the designer – who has a great understanding of the core technology – must effectively translate external demands (e.g. by executives) so that they harmonize with local conditions rather than cause disruptions. In order to accomplish this, he/she must be familiar with the terminologies (or syntaxes) of both worlds in order to conceptualize the current state as well as the desired future state. This is also true of information which adjusted to the actual department and individual work tasks gives individual advantages as well as rich interoperational flow (Fagerström, 2003, p.189). It is entirely possible that the core processes over time achieve a level of distinctiveness that is difficult to emulate or replace (McKiernan, 1997). Should this distinctiveness be considered valuable, it stands to reason that the organization would seek to preserve it – even if this prohibits closely knit organizational design.

To some extent, the bottom-up paradigm precludes strict objectivity. Where one individual may see a problem, another sees business as usual. This difference in perspective is a result of the intricate ways in which system, organization and context influence one another. Following this logic, defining the problem is no longer a technical issue, but a social one (Magoulas & Pessi, 1998, p. 130-132; Checkland, 2000). As technical and non-technical aspects of the organization continue to influence one another, unpredictable – emergent – properties appear over time.
A complementary view of the bottom-up perspective is given by Ciborra (2000, p. 26-27 who argues that separability of design and management is a misconception by management that holds no bearing on reality. Designers and operational staff exert a sense of care and cultivation in their work which enriches the core technologies. It is only then that alignment between human and artefact can be achieved (Monteiro, 2000, p.72-75).

Davenport & Prusak (1997, p. 177) argue that top-down approach is problematic where critical know-how is located at the operational level. Highly normative change efforts championed by top-level management are typically met with lukewarm enthusiasm and limited success. In these milieus, it is therefore wiser that management concern itself with coordination; for instance by identifying competencies, assigning responsibilities and clarifying the organization’s strategy and objectives. Davenport & Prusak (1997) sees this as the way to achieve real information interoperability rather than leave it at platitudes and ambitions. This strategy necessitates dealing with individuals in the organization who want to keep information to themselves for political, emotional or technological reasons. The information architecture therefore needs an appropriate level of inscription which will lead to behaviour that is beneficial to all parties (Monteiro, 2000, p.76-79).

5 Method
The aim of this study is to approach semantic interoperability from an informatics perspective by studying how it may be discussed in terms of human understanding. This is done by employing a qualitative research method; multiple case study (Yin, 2009, p.60-62). Since the study is undertaken in order to understand the efficacy of the framework, the study may be categorized as an explorative study. Our theoretical framework is tested against two organizations where primary data is collected via employee interviews. The analysis is conducted via a hermeneutical research process, which basically involves collecting and interpreting empirical data (Patton, 2002, p.113). As our analysis is intended to test theory, we therefore view this as a deductive analysis.

Theoretical framework was implicitly known to us before this study and further reading provided us with several frameworks and models relevant to our research. Since it was our intention to rely on extent literature in the field of Informatics we therefore selected theoretical models that easily could be augmented.

The hermeneutical process of interpreting case Beta started in an earlier study, where semantic interoperability was one of the benefits from structured information (Slumpi Persson, Ahlin & Öberg, 2011). Interest arose to understand general aspects of semantics and therefore empirical material from other company was needed. As the study was to be held on a general level and used to illustrate theoretical material, two companies, with different views, seemed appropriate. For the study at hand, one interview was held in each organization with personnel from middle management level – altogether two interviews. Sample size is in accordance with the fact that the empirical material will be used as illustrations and individual chosen due to their knowledge and experience. The interviews were held at the offices of the respective informants. We dedicated 45 minutes for company Alfa and 85 minutes for company Beta. The interviews were semi-structured, encompassing prepared- and unprepared questions based on informant responses (Creswell, 2007, p.352). The interviews were recorded and subsequently transcribed and categorized in different themes according to TFI-model. Some secondary data was also collected via the public websites of said organizations. Analysis was undertaken in an iterative fashion, where authors compare and discuss empirical material.
6 Case studies

In this section we present our case study and our results.

6.1 Case Alpha

Our first case is the branch office of a corporate group active in the financial sector. Altogether, the group encompasses approximately 15000 employees divided over 500 branch offices. Alpha offers a plethora of financial services to a wide range of customers. In order to provide customer service, the financial data of each client has to be accessible for every branch office. Given the nature of the information handled by the financial group, securing information from unauthorized access as while concurrently ensuring customer service are both primary concerns.

6.2 Case Beta

Our second case belongs to a global industrial group with approximately 12500 employees of which 2000 are working at Beta. Beta works with production and maintenance of electronic defence systems, mostly radar systems. Product development is usually based upon existing products, focusing on customization in correspondence with the wants and needs of the customer. Due to the nature of the products, large development projects involving several departments are standard operating procedure. In combination with product life-time service contracts, inter-departmental information is a necessity. External stakeholders such as subcontractors, auditors or users often require access to product documentation as a means to ensure quality.

6.3 Empirical pictures from cases

The centralized management style of case Alpha promotes semantic interoperability by means of uniformity – rules and regulations are formulated by top-level management and subsequently distributed through an intranet to national or global branches as relevant. It is the explicit duty of each employee to ensure that he/she keeps up with any and all changes in services and proper procedure. As for the comprehensibility of the information provided, our empirical data suggests the implicit assumption that education and on-the-job training is intended to ensure uniformity of interpretation. However, alterations in formal procedure are often absorbed via informal division of labour. Certain events considered to be of particular significance are sometimes brought up and discussed at local office meetings, thus providing an opportunity for further exposition.

Informal organizations consisting of different persons are set in place to keep us updated on changes in rules and regulations. Sometimes news is even brought up on office meetings.

(Senior manager, Alpha)

Alpha is characterized by strong centralization which also extends to the governance of information. Should a branch office wish to have something posted anywhere on the intranet, this would have to be communicated to the regional manager – and possibly further on up the hierarchy – for approval. In contrast to the management of the information structure, the actual information may in practical terms be considered the property of each office. Private customers as well as small businesses typically frequent the same branch office over a long period of time and it is not uncommon for business-owners to handle private- and business
finances at the same office. This serves as an impetus for inter-divisional work processes as the financial situation of an individual may depend just as much on how well his/her business is doing as the shape of his/her private finances. As the line between private- and business finances sometimes blur, the branch office staff must sometimes handle this complexity in an informal manner.

Case Beta provides a sharp contrast as top management of the corporate group adopts a more hands-off approach to the identities of its constituent companies. Each company has its own name, logo and product catalogue. Customer relations are also handled on a strictly individual basis - there is no shared customer registry. Hence, each company is left to manage its own information – limiting “global” information to the essentials such as financial data. However, within Beta, semantic interoperability is managed by means of a product revision standard. Originally implemented some 50 years ago, it was not primarily intended as a means to disseminate information, but rather safeguard documentation of complex artefacts. Beyond complexity, the lengthy life-span of the artefacts (in excess of 40 years) places high demands on the logic underlying documentation in order to trace variations in components and configuration over time, devoid of ambiguity. As this product revision standard has been at the core of the company’s product development for a long time, it is no longer a mere formalism. The codes and expressions stated in the standard are often used by employees in daily conversation. To some degree, this furthers the understanding of the artefact as a whole as the product revision standard is based on the type of product, its model, configuration et cetera. The logic underlying the product revision standard is managed by a small team within the company. Operating in an informal, democratic fashion, they maintain the standard, instruct new employees in its use and modify the standard when necessary.

No, I am the group leader, but we go by a democratic model. If three people [out of four] say OK, then it is OK. If two people say it is OK, then we have not achieved [the goal]. (System owner, Beta)

Inputting data is left to the individual employees who are tasked with modifying the design of the artefact in question. While mistakes do occasionally occur in applying the correct code to a revised design, this is rarely a problem as employees usually work within a limited range of possible variations to an artefact. The project manager is then ultimately responsible for ensuring that the artefact corresponds to customer specifications as well as documentation.

In comparing the two cases, there are obvious contextual dissimilarities – most noticeably the products and services that they provide. Yet they are similar in so far as they both depict organizations that have been in existence for over half a century – less than a decade of which spent under current ownership. It is also interesting to note that employees in both cases prefer to “ask a colleague” rather than utilize designated support functions.

7 Analysis/discussion

The discussion will follow the structure of the TFI-model with emphasis on relationships between the layers, as it is here that semantics, i.e. understanding, plays a significant role (Liebenau & Backhouse, 1990, p. 109-112).
7.1 Juxtaposition, Alpha and Beta

Illustrating the cases using the TFI-model, we may observe two distinct ways in which technology and informal organization may influence the formal aspects of the organization.

Figure 2: Alpha interpreted via TFI-model

Figure 3: Beta interpreted via TFI-model

The centralized information management at Alpha is motivated by a need for similarity despite serving customers nationwide. In the terminology of the semiotic framework, uniformity in pragmatics is sought by means of instructions and updates distributed via the intranet. A sense of shared semantics is attained on a local level when new instructions are contextualized by means of formal meetings or informal conversations. From a TFI perspective, we may describe case Alpha as being governed by formal directives distributed via technical (IT) systems and supported by means of informal structures that have formed in order to improve contextualization.
The specialized information management at Beta stems from its need to trace variations in complex products over a very long time-frame. The elaborate syntax set in place to facilitate this necessity has over several decades become ingrained in the organization to the point where it is practically a part of local culture. Although initially confronted with a steep learning curve, employees are able to use the highly codified syntax to share information in a very efficient manner; supported by a computerized catalogue-system if it is needed. Hence, we view the information infrastructure at Beta as being governed from cultural, informal structures present in the organization with technical systems merely serving a supporting role.

7.2 Technical systems and formal organization

In reviewing the two cases, we can discern two distinct ways of managing technical systems. Alpha resembles what Davenport & Prusak (1997, p. 74-75) describe as a monarchy; the ability to alter technical systems is highly centralized. This enables uniformity in data quality and the ability to handle customers the same way despite serving a vast geographical area and heterogeneity in customer requirements. Beta corresponds to feudalism where cooperation, integration and technical conformity between group companies is limited. This allows great disparity in products, services and customer base.

In terms of syntax, it is readily apparent that new employees face different levels of complexity in the two cases. Workers at Alpha undergo on-the-job training that is usually completed within the space of one day. Employees at Beta face a more drawn out period of intermittent training that covers several months. Based on the disparity in time-frames, we may infer that the level of generality between syntaxes differ significantly, Beta being far more idiosyncratic. Furthermore, we may discern different levels of inscription in the technical systems. At Alpha, employees are to a large extent instructed to act based upon information that is stored in the technical systems. In effect, the technical systems act as a means to impact the actions and understanding of employees from a single point in the organization. Beta provides a sharp contrast in that information stored in the system provides a distinct logic for arranging information, yet does not specify behaviour in any explicit form. This is most acutely felt through the lack of regulatory functions despite the complicated syntax. Employees are in other words “free” to make mistakes.

7.3 Formal and informal organization

As previously mentioned, the levels of complexity facing new employees differ significantly between the two cases. Alpha sees new employees able to learn to utilize the technical systems within days whereas the product configuration system at Beta takes months to learn. This difference is noteworthy seeing as both companies typically hire college/university graduates with relatively homogeneous backgrounds. The semantic structure (Langefors, 1973, p. 242-249) of employees should therefore be somewhat similar with respect to skill-set. Even so, Beta accepts a lengthy process of adaption whereas Alpha has apparently taken steps to ensure that pre-existing skills (attained at university) to a large degree suffice. The apparent acceptance of this lengthy time frame suggests a high level of integration into corporate culture (Monteiro, 2000, p.72-75).

Moving beyond the lengthy period of adaptation, the common language provided by the product revision system at Beta allows the organization to economize on information (Boisot, 1995, p.39-86) by simply using product codes as short-hand. Also, employees are rewarded for their efforts by having a large extent of freedom to work as they see fit. The situation at Alpha is quite the opposite as information is made available in greater volumes than
employees are able to process; effectively bringing about scanning or informal division of functional areas.

8 Conclusion

The aim of this study was to approach the topic of semantic interoperability from an informatics perspective, i.e. in terms of human understanding rather than technical compatibility. A literary study suggested two extreme viewpoints in managing semantic interoperability: Top-down and bottom-up. Utilizing a deductive approach, a case study has been undertaken in order to illustrate what form these viewpoints may assume in real life.

Our real-world cases have been analysed based upon a view of semantic interoperability based upon semiotics, the study of signs. Furthermore, our view of the cases is based upon the TFI-model which expresses three organizational layers: Informal organization, formal organization and technical systems. Combined, the human focus of semiotics and the explicit consideration of formal as well as informal factors of TFI provide us with a novel perspective on semantic interoperability. We believe this perspective is pertinent to academia as well as practice due to its generality. However, we do not claim this to be a useful tool as it stands, but merely a suggested starting point for future endeavours.

Approaching the concept of semantic interoperability by means of a common body of research would enable a wider research community than each individual context can produce. We therefore call on a unified approach so that we may all learn from one another rather than suffer from a lack of semantic interoperability ourselves.

9 References


