A Vector-Based, Content-Analytic Methodology for Comparing Negotiated IT Service Level Agreements

Daniel S. Soper
Department of Information Systems, W.P. Carey School of Business, Arizona State University
Daniel.Soper@asu.edu

Michael Goul
Department of Information Systems, W.P. Carey School of Business, Arizona State University
Michael.Goul@asu.edu

Haluk Demirkan
Department of Information Systems, W.P. Carey School of Business, Arizona State University
Haluk.Demirkan@asu.edu

Eileen Aranda
Department of Supply Chain Management, W.P. Carey School of Business, Arizona State University
Eileen.Aranda@asu.edu

Luis Aranda
Department of Supply Chain Management, W.P. Carey School of Business, Arizona State University
Luis.Aranda@asu.edu

ABSTRACT

Growth in the outsourcing of IT services has led many organizations to enter into negotiated contractual agreements with both internal and external service providers known as information technology service level agreements (IT SLAs). To further empirical research into IT SLAs, we present a methodological approach based on the theory of conceptual spaces that allows the content of these negotiated agreements to be analyzed and compared geometrically using vector representation. The outcome of such an analysis is a set of distance measures by which defensible statements regarding the similarity of IT SLAs can be made. We also discuss how the comparisons provide insights for IT SLA negotiation researchers that go beyond alternate empirical and analytic methods. The implications of these comparisons are discussed with a specific focus on research methods that can foster cumulative empirical investigation of IT SLA negotiation support system requirements.

Keywords

Service level agreements, outsourcing, negotiation support, conceptual spaces.

INTRODUCTION

The increasing utilization of IT outsourcing as an organizational business strategy has fostered a renewed interest in understanding the nature of the contractual agreements that bind IT service providers and service requesters. In addition, the processes by which those contracts are negotiated, formalized, and maintained are becoming increasingly important to organizations' bottom lines. These contracts, known as information technology service level agreements (IT SLAs), specify the relationship between parties by rigorously defining the IT services to be provided to a purchasing organization and the performance metrics by which the delivery of those services will be evaluated over time. As contractually-defined IT outsourcing relationships become both more common and more complex, the means by which those relationships can be aligned with organizational business strategies will become an increasingly important area for information systems research.

Given the proliferation of IT SLAs in outsourcing relationships, it is critically important that researchers have access to robust research methodologies by which those contracts and interorganizational relationships can be effectively 'examined'. To examine such contracts implies that one needs a frame of reference for performing contract comparisons, abstracting definable and fundamental SLA properties and a capability to generalize SLA metrics across varying domains within which they are negotiated. The dynamic nature of IT SLA relationships, coupled with the innately complex and domain-specific structure of IT SLA contracts, has made it problematic to generate investigative inroads into this research domain. Given this intrinsic complexity, very few methodological tools currently exist that are well-suited for facilitating useful research in this area (Demirkan, Goul and Soper, 2005). To that end, the development of efficacious and defensible methodological tools is of high priority to researchers working in this field.

One of the principal impediments to understanding the processes and mechanisms underlying IT outsourcing relationships is the lack of a methodology by which the contracts that define those relationships can be evaluated and compared. As a means of addressing this problem, the current paper develops and presents a distance-based methodological approach grounded in the theory of conceptual spaces that is intended to allow for the effective comparison of the IT SLA artifacts. It is expected that this methodological approach will serve to inform and resolve many IT SLA-based research questions, thereby yielding an enhanced understanding of the processes active in the IT outsourcing domain.

The balance of this paper is organized as follows: Section 2 discusses the geometric spatial structures and theory that together form the foundation of the methodological approach. Section 3 outlines the methodology itself and provides an overview of its key components. Finally, a brief summary and several concluding remarks are provided in Section 4.

THEORETICAL BACKGROUND

One of the most fundamental problems impeding IT SLA research is the difficulty that exists in comparing seemingly different SLA artifacts in a statistically sound way. From an investigative standpoint, researchers require a valid means of assessing the degree of similarity between a set of contracts if any inferences are to be made about the processes underlying their generation and maintenance; *i.e.* a representational framework is needed by which the content of IT SLAs can be described and compared. Given that this problem is rooted in similarity assessments, geometric spaces and structures afford a logical foundation for a representational framework, as they allow similarity relationships to be modeled in a natural way (Gärdenfors, 2000).

Employing geometric spatial structures as a representational framework is by no means new; researchers in many fields have used geometric spaces to impose structure on similarity problems. In cognitive psychology, for example, geometric spaces have been used to examine how and when children develop dimensional reasoning capabilities (Carey, 1985; Goldstone and Barsalou, 1998), as well as how people psychologically represent and interpret color (Gallistel, 1990). Geometric representations have also been used successfully in linguistics to explore the categorical perception of phonological units (Petitot, 1989) and cognitive grammar (Langacker, 1991). These structures have further been employed in artificial intelligence to examine artificial vision (Chella, Frixione and Gaglio, 1997) and the processing of activation spaces (Churchland and Sejnowski, 1992), and in organizational science research to derive implementation process taxonomies (Sabherwal and Robey, 1993). The examples provided here are but a few of many; spatial representations have been used in several additional fields including sociology, neuroscience, biology, economics, and philosophy. Clearly, the utilization of geometric spaces as a representational framework is not restricted to any particular domain, as the concept has been widely applied with a great deal of success to myriad disparate research areas across many disciplines.

The theory of conceptual spaces was developed as a spatial framework for representing information on a conceptual level (Gärdenfors, 2000). When this representational framework is combined with a geometric structure and a series of elemental connections postulated for a specific domain, the result is a set of empirically testable propositions by which scientific knowledge of the target domain can be advanced. At its core, the theory of conceptual spaces allows for the description of objects within a multidimensional geometric space whose dimensions are defined by the domain of interest. Because every object within a conceptual space has a location that can be characterized in terms of the dimensions that describe the space, the theory allows for the computation of similarity judgments as a function of the geometric distance between two objects; *i.e.* as the distance between two objects in the conceptual space decreases, those objects can be said to be growing increasingly similar. The theory of conceptual spaces therefore provides researchers with a natural framework by which the degree of similarity between conceptual objects can be represented.

Dimensions are used within the theory of conceptual spaces to describe the nature of a geometric space, thereby providing a measurement model by which object locations can be specified. The dimensions that comprise the geometric structure of a given problem domain must therefore be carefully selected by the researcher given the context of the phenomenon under investigation. The following example should serve to clarify this notion: Suppose that an information systems researcher wishes to determine the degree of similarity between several IT SLAs. For the sake of simplicity, presume that she desires her similarity judgment to be based exclusively on two criteria: (1) the level of hardware support specified in the contract, and (2) the cost of the contract. Within the boundaries of her investigative domain, the cost and hardware support properties serve as the two evaluative dimensions. For any set of IT SLAs, the degree of similarity between contracts can be determined by plotting their locations geometrically in a space defined by the evaluative dimensions, and then calculating the distance between those locations. This process is illustrated for a set of three IT SLAs in Figure 1.

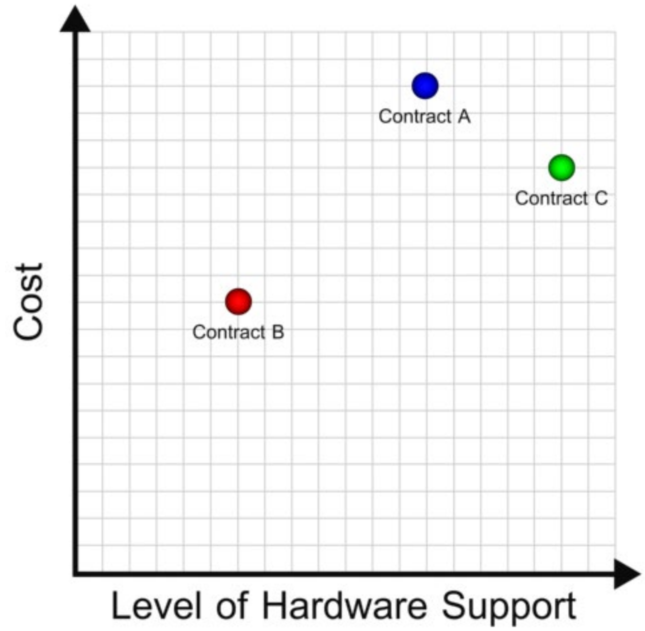

Figure 1. Two-dimensional space for IT SLA comparisons

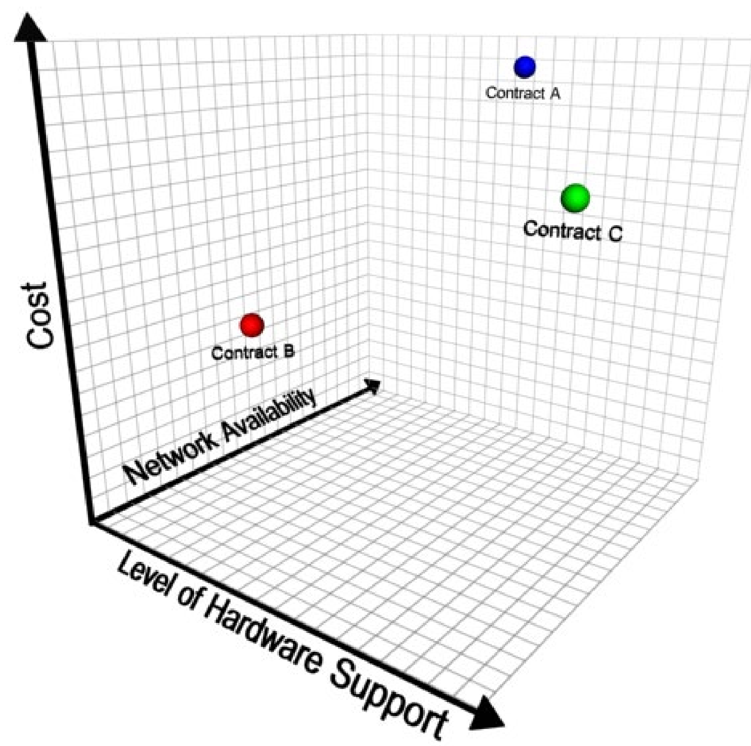

Figure 2. Three-dimensional space for IT SLA comparisons

Given the definition of the conceptual space as a function of two dimensions (*i.e.* IT SLA cost and level of hardware support), Contracts A and C in the figure are clearly more similar to each other than either is to Contract B. This similarity judgment is based on an assessment of the geometric distance between the points that define each contract's location in the conceptual space. Note that if an IT SLA does not contain a particular dimension, it can still be legitimately compared to other IT SLAs within the conceptual space. If, for example, Contract B in Figure 1 did not contain any provisions for hardware support whatsoever, the contract could still be plotted in the conceptual space simply by assigning it a value of zero for that evaluative dimension. More generally, the upper and lower bounds of any dimension within a conceptual space can be defined by the observed properties of the set of objects being compared.

The true power of this representational approach becomes readily apparent when one extends the geometric structure by adding more evaluative dimensions to the conceptual space. Figure 2, for example, extends the IT SLA comparative framework by adding a third dimension representing network availability. The degree of similarity between the IT SLAs can clearly still be assessed as a function of the geometric distance between them in the three dimensional space. Although human cognitive and perceptual limitations tend to prohibit us from envisioning more than three simultaneous geometric dimensions, a conceptual space can be mathematically extended into as many dimensions as are necessary to fully describe the domain of interest.

After the dimensions of interest have been defined, it becomes possible for any IT SLA to be fully described within a multidimensional spatial domain simply by examining its geometric *location* within the space. Furthermore, if the order of property descriptions and the metrics by which those properties were measured are preserved, the description of an IT SLA in the geometric space can be reduced to a vector. In the scenario depicted in Figure 2, for example, Contract A has a cost of 19 units, a level of hardware support of 10 units, and specifies 18 units of network availability. If it is known that the properties of an IT SLA will always be listed in the same order (*e.g.* cost, followed by hardware support, followed by network availability), then a vector of numbers can be used to meaningfully describe *any* IT SLA within the space. The vector description of Contract A as specified in Figure 2, for example, would be [19, 10, 18]. It is important to note that the scales by which the evaluative dimensions are measured need to be standardized before any distance calculations are attempted if the researcher wishes to avoid the problems associated with assigning dimensional weights.

Given the vector-based approach illustrated above for describing SLA locations, and presuming (1) that the dimensional scales have been standardized, and (2) that the primitive geometric relations of equidistance and betweenness hold for the conceptual space (*i.e.* the space can be formally designated as a metric space (Weisstein, 2005)), it becomes possible to mathematically calculate the Euclidean distance d between any two SLAs x and y in an n-dimensional space by utilizing the following formula:

$$d_{(x,\,y)} = \sum_{i=1}^{n} \sqrt{(x_i - y_i)^2}$$

The Euclidean metric is ideal for similarity assessments in conceptual spaces, as the distance between two objects in a Euclidean space remains constant regardless of how the dimensional axes defining the space are rotated (Gray, 1997). By calculating the distance between two IT SLAs, a researcher is simultaneously arriving at a metric by which the similarity between those SLAs can be assessed; *i.e.* similarity judgments and geometric distances are intimately related in a conceptual space. The nature of the functional relationship between similarity and distance is widely agreed to be an exponential decay function, such that the similarity between two IT SLAs decays exponentially as a function of the distance between them. Two expressions have been put forth to formulaically define this relationship. Given two SLAs x and y, the distance between those SLAs $d_{(x,\,y)}$, and a sensitivity parameter c, the first function (which takes the form of a Gaussian function (Nosofsky, 1986)) defines the similarity s between the SLAs as:

$$s = e^{-c d_{(x,y)}^{\,2}}$$

Other researchers favor the second function (Shepard, 1987), which defines the similarity between the SLAs as:

$$s = e^{-c d_{(x,y)}}$$

Regardless of the differences between these two functional expressions, similarity is widely agreed to be an exponentially decaying function of distance as shown in Figure 3.

Although heretofore only spatial dimensions have been considered, it is also important to understand the interaction of space and time within this framework. As time passes, the location of an IT SLA within a geometric space is likely to change in response to the shifting IT needs of the purchasing organization, or in response to changes in the supplying organization's ability to provide services. These changes, which occur during contract renegotiation cycles, can be tracked using conceptual spaces simply by taking repeated measures of the SLA's properties at the appropriate points in time. If the series of measures are plotted in a geometric space, the researcher can then analyze the *path* of the SLA across time in an effort to better understand the processes and trends underlying contract-based IT outsourcing.

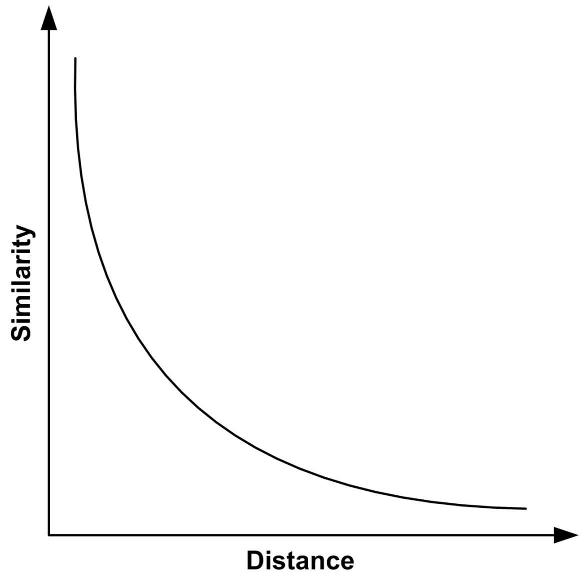

Figure 3. Exponential decay relationship between Distance and Similarity

Perhaps the most powerful attribute of the representational framework put forth by the theory of conceptual spaces is that it allows a researcher to specify a geometric structure describing a particular domain, and translate the objects, properties, locations, and movements within that domain into a set of empirically testable hypotheses. The theory of conceptual spaces can further be used to facilitate the construction of new theoretical frameworks, as a researcher needs only to specify the connections between seemingly disparate geometric domain structures in order to arrive at a representation of a new theory. A theoretical representation specified in this way has the advantage of allowing for the formulation of new functional laws through inductive inference, thereby making possible more refined and informed theoretical predictions (Gärdenfors, 2000).

METHODOLOGICAL OVERVIEW

The methodology described herein proposes that the content of IT SLA contracts can be described by multidimensional vectors, and that the similarity between any set of contracts can therefore be determined as a function of the distance between the spatial locations described by those vectors. This approach allows for quantitative similarity judgments to be made between and among negotiated contracts at multiple time periods. It is hoped that this technique will facilitate research that informs not only knowledge about the mechanisms underlying IT SLA negotiation, but more broadly knowledge regarding the processes, trends, and patterns underlying IT outsourcing in general.

In translating the theory of conceptual spaces into a methodological approach for analyzing IT SLA content, it is first necessary to develop the dimensions that will define the geometric space by which the contracts of interest can be described. Identification of the dimensions that define the concept space requires that an initial analysis of the content of all of the IT SLA artifacts under investigation be performed with the intent of isolating all of the service categories contained therein. Several approaches to this content analysis are possible. In one approach, several trained judges perform the identification task independently on all contract artifacts, after which disparities in the identified service categories and category descriptions are reconciled. This approach yields enhanced validity by minimizing any bias introduced during the categorization process. Alternatively, if an existing validated categorical structure can be identified that can be shown to fully subsume the content of the contract documents, that structure can be used in lieu of the approach just described. Another approach is to use text analysis software to identify dimensional clusters in the set of contracts. The researcher then simply examines the clusters, and assigns each an appropriate name. After completing this identification task, the researcher will possess a list of IT services that constitute the *service set* for the IT SLAs. Note that the mathematical model described previously allows for contracts to be comprised of distinctly different service categories; *i.e.* any given IT SLA contract can be comprised of any subset of the categories contained in the service set. There is no requirement that all of the contracts share the same set of service categories, although it is expected that significant overlaps among categories will exist between IT SLA contracts.

Each of the categories in the service set represents a single dimension in the conceptual space. Although there is no upper limit restricting the number of dimensions that can be contained in the geometric representation framework, it should be explicitly understood that the concept space for the target domain is contextually relevant only to the categories contained in the current service set. If a category is added to (or subtracted from) the service set, the underlying geometric structure of the representation framework is changed, as is the nature of the target domain itself. In these circumstances, any similarity judgments obtained from analyses conducted using the prior representation framework cannot be assumed to be valid within the spatial structure that defines the new target domain.

The most challenging task associated with this methodological approach lies in specifying the multidimensional vectors that represent each of the IT SLA contracts in the Euclidean *n*-space. The goal of this task is to classify the content of each of the contract documents by the categories contained in the service set. From a high-level perspective, the researcher assigns a non-negative value weight to each category in the service set by which the relative significance of the category to the current IT SLA can be communicated. In practice, a substantial problem exists in this task with respect to measurement. As it is classically treated, measurement is "the assignment of numerals to objects or events according to rules" (Stevens, 1951). Assigning appropriate and valid value weights therefore requires that a set of rules be developed by which the relative significance of the categorical content of the IT SLAs can be consistently and objectively evaluated.

Regardless of the content analytic technique ultimately chosen by the researcher, there are several general guidelines that should be adhered to when analyzing IT SLA contracts. Foremost among these principles is the utilization of a ratio scale for the classification task. Ratio scales are desirable for the classification task because they are widely acknowledged to utilize the most sophisticated level of measurement. The employment of a ratio scale requires not only that the content of the IT SLAs be classified into orthogonal categories using numbers that have a meaningful zero point, but also that the differences between those numbers be equal so that any two values can be expressed as a valid ratio (Neuendorf, 2002). When analyzing the content of IT SLAs along a service dimension, a value of zero should indicate that the contract does not address the current service dimension at all, while a number greater than zero should indicate the contextual significance of the current service to the contract under examination *relative to the other IT services addressed therein*.

While it is beyond the scope of the current paper to debate the merits of the myriad approaches that have been developed to address the textual content analysis problem, it is possible to address the two specific approaches that are most applicable and valid for IT SLA negotiation: (1) a manual approach in which trained judges catalog the content of the contracts by hand, and (2) an automated approach in which a software agent performs the analysis. Each of these approaches has distinct advantages, and the choice of which to use must be left to the researcher. Before using either approach, however, the researcher must construct a set of rules defining how the content of the contracts will be categorized along the previously identified spatial dimensions. If the manual approach is chosen, the analytic rules may be as simple as a formal set of definitions for each category by which the judges can evaluate the contracts, or may be as complicated as highly-detailed codebook. The goal of the researcher in employing a rule set in the manual approach is to eliminate individual differences among judges, thereby increasing the validity of the research. If the automated approach is chosen, however, the most common approach to defining analytic rules is to employ a series of 'dictionaries' by which the text analysis can be carried out. As used here, a dictionary is "a set of words, phrases, parts of speech, or other word-based indicators that is used as a basis for a search of texts" (Neuendorf, 2002). For automated IT SLA contract analysis, a single dictionary should be developed for *each* spatial dimension in the service set. The content of each contract is then analyzed using *every* dictionary during the automated classification process.

If the manual content analytic approach is chosen for the classification task, the researcher should ideally train a series of independent judges, each of whom will analyze the full set of contract documents independently with respect to the spatial dimensions and associated rules defined for the domain. Bias can be reduced by ensuring that the judges work independently, and are unaware of the intent of the research. Upon completing their task, the interrater reliability of the judges' independent assessments must be determined in order to ensure the validity of their classifications. A minimum interrater reliability of 0.80 is widely considered necessary for validity purposes.

If the automated content analytic approach is chosen for the classification task, the researcher must carefully select a text analysis software program that will effectively meet his or her needs. Fortunately, there are dozens of commercial text analysis programs to choose from, one of which is likely to meet the needs of the researcher, and almost all of which generate the frequency output necessary for vector construction. Alternatively, an ambitious researcher could develop a custom automated text analysis agent specifically engineered for the task at hand. The advantage of selecting a commercially available text analysis product is that the tool itself carries a sizeable portion of the responsibility for the reliability and validity of the results, whereas with a custom product the burden of responsibility falls squarely on the shoulders of the researcher. Additionally, most commercial products are not forthcoming with the means by which they arrive at their results. This problem is not an issue with a custom software agent, as the researcher has full control over the internal algorithmic processes by which the results are produced.

The result of the classification task -- regardless of the analytic approach taken -- is a set of value weights describing the significance of each category to every contract document. Following classification, the researcher must establish a fixed order for the dimensions so that they can be described as vectors. The specific order chosen is of no consequence, presuming that the same order is applied to every artifact. The value weights for each category must then be standardized across all of the IT SLA contracts in order to ensure dimensional equidistance when the standardized value weights are plotted as vectors in the geometric space. The outcome of these ordering and standardization processes is a single multidimensional vector for each IT SLA contract that represents the content of the contract with respect to the categories contained in the service set.

Given the set of artifact vectors that result from the content analysis process, plotting each IT SLA in the domain-specific geometric space becomes a straightforward task. Each vector simultaneously represents the content of its associated IT SLA contract *and* a mathematically definable location in a multidimensional geometric space. Because the spatial framework underlying the vector locations is constant across the full set of IT SLAs under examination, the locations of all of the contracts can be simultaneously plotted in the geometric space in a meaningful fashion. As described previously, this common geometric representational framework allows distance measures to be calculated by which the researcher can make valid and defensible statements regarding the similarity between and among the set of IT SLA artifacts.

The methodological approach developed thus far is limited in that it only allows a researcher the ability to examine and compare a set of IT SLA artifacts at a fixed point in time. Although this cross-sectional approach is valuable in its own right, the methodology can be easily extended into the realm of space-time. The incorporation of a time dimension into the spatial coordinate system vastly extends the power of the methodological approach by allowing for longitudinal assessments. Rather than assessing the similarities between different IT SLA contracts, incorporating a time dimension could, for example, allow a researcher to examine how a single IT SLA changes as it is renegotiated over time. The knowledge gained from such an analysis would allow the researcher to make inferences about an organization's changing IT needs and associated strategies. Following past research (Demirkan, et al., 2005), this approach could also be used to inform our understanding of the IT SLA negotiation process by allowing preliminary IT SLA artifacts prepared by the separate parties in an IT outsourcing relationship to be compared to the final negotiated IT SLA contract. As is evident in these examples, changes in vector representations of contract content over time can serve to further understanding of the processes underlying the creation and maintenance of those IT SLA contracts. Process knowledge of this sort will be critically important in informing the design of future negotiation support systems.

SUMMARY AND CONCLUDING REMARKS

This paper presented a methodological approach built on geometric spaces and distances that can allow for informed judgments to be made regarding the similarity of IT SLAs. These agreements, which result from formal negotiation processes between service requesters and service providers, are critical in furthering our understanding of IT outsourcing relationships. Given the growing utilization of IT outsourcing as an operational business strategy, an understanding of these contractually-defined relationships will become an increasingly important component of managerial information systems research. It is hoped that the vector-based approach presented herein will facilitate additional research not only into the role of IT SLAs in the outsourcing relationship, but also into the negotiation processes by which those agreements are generated and maintained. Ultimately, research conducted using methodological tools such as the one described by the current paper may inform the design of automated systems that can support the negotiation of IT SLAs between service requesters and service providers. Negotiation support systems built on solid process knowledge have the potential to provide valuable competitive advantage to organizations negotiating the outsourcing of their core IT services.

The importance of developing a suitable methodology for conducting research in IT SLA negotiation is also significant from a business strategy standpoint. Outsourcing organizations need such methodologies to provide a means to compare alternate SLA artifacts proposed by competing vendors. In addition, service providers need a means to evaluate contracts under negotiation to assess their capability to perform and to determine if what they are promising to deliver is significantly different than the service levels negotiated with other companies. At stake is the ability to maintain a uniform market presence so that one customer does not perceive that another customer achieved a significantly 'better' contract for the same price. In addition, if an organization has a portfolio of outsourcing contracts, a suitable methodology -- such as that proposed herein -- may well be a step towards identifying risk measures based on the criticality of business processes and the associated provider network's agreed-upon ability to perform.

In conclusion, this paper has described a validated methodology for conducting research into IT SLA negotiation. It was argued that the methodology's validity to IT SLA negotiation is substantiated due to its historical usefulness in similar domains of study where the approach has been shown to be effective in leading to a cumulative research agenda. In future work, we plan to further validate the methodology in a pilot study and later in a more comprehensive field study. Our intent here has been to make a contribution to the ongoing research design and methodology dialogue in the increasingly important domain of IT SLA negotiations.

REFERENCES

Carey, S. (1985) Conceptual Change in Childhood, MIT Press, Cambridge, MA.

Chella, A., Frixione, M. and Gaglio, S. (1997) A Cognitive Architecture for Artificial Vision, *Artificial Intelligence,* 89, 73-111.

Churchland, P. and Sejnowski, T. (1992) The Computational Brain, MIT Press, Cambridge, MA.

Demirkan, H., Goul, H. and Soper, D.S. (2005) Service Level Agreement Negotiation: A Theory-based Exploratory Study as a Starting Point for Identifying Negotiation Support System Requirements, *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, Kona, HI.

Gallistel, C. (1990) The Organization of Learning, MIT Press, Cambridge, MA.

Gärdenfors, P. (2000) Conceptual Spaces: The Geometry of Thought, MIT Press, Cambridge, MA.

Goldstone, R. and Barsalou, L. (1998) Reuniting Perception and Conception, *Cognition,* 65, 231-262.

Gray, A. (1997) Euclidean Spaces, In *Modern Differential Geometry of Curves and Surfaces with Mathematica*, CRC Press, Boca Raton, FL.

Langacker, R. (1991) Foundations of Cognitive Grammar, Stanford University Press, Stanford, CA.

Neuendorf, K. (2002) The Content Analysis Guidebook, Sage, Thousand Oaks, CA.

Nosofsky, R. (1986) Similarity, Frequency, and Category Representations, *Journal of Experimental Psychology,* 115, 39-57.

Petitot, J. (1989) Morphodynamics and the Categorical Perception of Phonological Unit, *Theoretical Linguistics,* 15, 25-71.

Sabherwal, R. and Robey, D. (1993) An Empirical Taxonomy of Implementation Processes Based on Sequences of Events in Information System Development, *Organization Science,* 4,4, 548–576.

Shepard, R. (1987) Toward a Universal Law of Generalization for Psychological Science, *Science,* 237, 1317-1323.

Stevens, S. (1951) Mathematics, Measurement, and Psychophysics, In *Handbook of Experimental Psychology*, Wiley, New York, NY.

Weisstein, E. (2005) Metric Space, In *MathWorld*, Wolfram Research, Inc., Champaign, IL.