Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2010 Proceedings

International Conference on Information Systems (ICIS)

2010

ONTOLOGY-BASED INFORMATION EXTRACTION FOR ANALYZING IT SERVICES

Amit V. Deokar
Dakota State University, Amit.Deokar@dsu.edu

Sagnika Sen
Penn State University, sus45@psu.edu

ONTOLOGY-BASED INFORMATION EXTRACTION FOR ANALYZING IT SERVICES

Research-in-Progress

Amit V. Deokar
Dakota State University
Madison, SD 57042
amit.deokar@dsu.edu

Sagnika Sen
Penn State University
Malvern, PA 19355
sus45@psu.edu

Abstract

Service Level Agreements (SLA) for multi-service Information Technology (IT) outsourcing contracts contain vast amounts of textual information. The SLAs provide details about a specific service, Key Performance Indicators (KPI) to measure its performance; as well as process elements, such as activities, events, and resources that are integral in achieving performance goals. However, KPIs and the process elements may be interrelated. The knowledge of such interrelationships is often tacitly present in the SLAs. The aim of our research is to extract this hidden information from IT service contracts and analyze them to empower customers of IT services to make better performance management and incentive decisions. We apply an Ontology-Based Information Extraction (OBIE) approach in developing a prototype decision support framework, named SLA-Miner. The results, obtained from analyzing a set of Industry SLAs, demonstrate the utility of SLA-Miner in identifying KPI interrelationships, deficiencies, and impacts of various process elements on individual KPIs.

Keywords: Information Technology Service Management (ITSM), Performance Management, Service Level Agreement (SLA), Ontology-Based Information Extraction (OBIE), Business Process Management

Introduction

Organizations are increasingly outsourcing their Information Technology (IT) infrastructure to third party service providers. This trend is growing fast with the advent of new service models such as SaaS (Software as a Service), cloud computing, etc. With such a growth of different outsourcing models, it is imperative that organizations invest in effective measurement and monitoring of their IT service contracts (KPMG 2007). "*The sheer volume of data available, coupled with the critical nature of effective decision-making by senior leadership teams, demands that large IT organizations declare the IT measurement program as a core competency*"(Cleary 2008). For large multi-service IT outsourcing contracts, performance management systems play a crucial role in incentive design and subsequent compensation scheme (Fitoussi et al. 2008). However, quite a few challenges exist in implementing and realizing the benefit of such performance management programs. One such challenge arises from managing the contracts in the form of Service Level Agreements (SLAs) (These SLAs can be executed between two different organizations or between the IT division and other business divisions within the same company). A typical IT outsourcing arrangement consists of multiple services being outsourced (e.g. network, office applications, e-mail etc.). Such a contract consists of multiple SLAs, typically one for each service. Each SLA includes detailed service descriptions, performance goals, (multiple) Key Performance Indicators (KPIs), penalty and reward clauses etc. This results in lengthy and complex set of SLAs often referred to as "*suicide pacts*" (Taylor et al. 2006) that render any measuring and monitoring difficult, if not impossible.

In this context, our research is aimed at identifying the interrelationships among KPIs and different process elements – such as activities, resources, and events - from the SLA documents representing an IT infrastructure service bundle. Our focus on interrelationships is driven by the fact that the different activities/resources/events required in IT services are highly interdependent. For example, lower bandwidth may impact the speed of query retrieval, even though network and data management may be treated as two different services in the contracts. Similarly, improper hardware configuration might lead to software malfunction. Since the same activities/resources/events may affect multiple performance measures (KPIs), it is important that their relationship is properly understood; especially by the customer organization in order to be able to align mutual interests through appropriate incentives. Furthermore, it is possible that some interrelationships are not apparent even to the service provider. Systematically eliciting such information will help in better resource planning and service delivery options.

In this paper, we propose an Ontology-Based Information Extraction (Chen-Yu et al. 2005; Cunningham et al. 2006; Li et al. 2007b) framework to identify interrelationships among the KPIs and the various process elements in IT services. This framework is instantiated in a prototype named SLA-Miner. By their very nature, SLA documents contain formal and semi-formal knowledge regarding IT service processes (e.g. *Network Services* will talk about *network speed/bandwidth* as a KPI, and *Database Services* will definitely mention *Transfer Performance* as a KPI). Hence an ontological approach seems appropriate as the implicit structural knowledge can be utilized to extract mutual dependencies that are not immediately apparent (e.g. any mention of *network speed/bandwidth* in the *Database Services* implies that the latter may have some level of dependency on the former). In other words, we take a sense-making approach where (semi)-codified knowledge is utilized to identify non-codified and tacit knowledge.

As such, our work contributes to the stream of research in IT services from two perspectives – IT Service Management (ITSM) and Service Computing. From an ITSM standpoint, we identify process and KPI interrelationships, with the main aim being the alignment of incentives with performance measures. That is, our focus on intelligence gathering at the process level (activities/events/resources) helps in operational workflow management of services. From a Service Computing point of view, the proposed OBIE framework (detailed later in the paper) may be implemented as the managerial layer that would help organizations in the continuous management and improvement of service contracts.

We utilize a set of real-life SLAs for a multi-year multi-billion dollar contract between a large government organization and a prominent service provider as a proof-of-concept for our proposed framework. Detailed description of the design and development of the framework, along with preliminary results are provided in the following sections.

Related Work

Interrelationships among service processes pose a significant challenge in the performance management of IT Services (Hayes 2005; Munk 2006). While it is acknowledged that complications arise in both service delivery and performance monitoring from interrelated processes, so far there is no systematic way to capture and model such interrelationships. While industry initiatives such as ITIL (IT Infrastructure Library) and COBIT (Control Objectives for Information and related Technology) are significantly influencing the practice of IT services management to look beyond the 'silo' orientation to IT infrastructure and take a process-oriented view, neither of these frameworks present a methodology to address the issue of interdependency. Some academic research has investigated the impact of process interdependence on contract performance (Mani et al. 2006), coordination patterns (Edgington et al. 2010), automated representation of knowledge concepts in SLA Management (Paschke et al. 2008). Overall, this remains an underrepresented area in academic research.

On the other hand, there has been significant effort in eliciting process level information from various organizational records. A notable effort is in the area of process mining which aims at analyzing process, control, data, etc. based on event logs produced by various information system applications (van der Aalst 2005), or applying data mining techniques on performance indicators as a means of identifying problematic areas in business processes (Grigori et al. 2004).

While these methods provide a conceptual direction to our work, they are not directly applicable in the present context, as our aim is to gain knowledge of process interdependencies *before* service delivery. Consequently, event logs and actual performance measures might not be available and/or applicable. Since Service Level Agreements contain enormous textual information, we adopt a relatively new approach (following Li et al. (2007a; 2009)), where process level information can be extracted from policy level documents (SLAs in our context).

Information Extraction (IE)

In developing a framework to extract information from service contract documents for identification and analysis of service interdependencies, we leverage information extraction techniques studied as part of natural language processing (Appelt 1999; Cowie et al. 1996; Cunningham et al. 2006; Pazienza 2003). The Message Understanding Conference (MUC) program (MUC-7 1998) in the late 1990s defined IE to consist of five main tasks. First, named entity recognition (NE) is concerned with identifying and classifying entity information (e.g., service names, activities). Second, coreference resolution (CO) is concerned with identifying different occurrences of the same entity, involving both anaphoric resolution (e.g., 'that' referring to 'help desk' in a sentence) and proper-noun resolution (e.g., different spellings for the same entity). Third, template element production (TE) builds on previous tasks to associate descriptive information to entities (e.g., finding different aliases for service names from the text). Fourth, template relation production (TR) refers to identifying small relations between entities (e.g., a service quality associated with a service). Fifth, scenario template extraction (ST), which is a complex task, refers to tying together information from TE and TR to infer complex analytic relations (e.g., a service is related to another service through a specific activity). The more recent Automatic Content Extraction (ACE) conceptualization (ACE 2000-2005) details three main objectives of IE: entity detection and tracking (EDT), relation detection and characterization (RDC), and event detection and characterization (EDC). These objectives can be mapped to the earlier MUC-7 conceptualization.

The process of accomplishing these IE tasks or objectives follows two main approaches: 1) knowledge-based approaches utilize a pre-defined conceptual representation of a domain of interest, whereas 2) machine learning/statistical approaches use large volumes of data to elicit knowledge elements in a given domain. Learning-based approaches require training data, which needs to be either readily available or created manually or semi-automatically. Currently, definitions about the various components of IT services (e.g. service descriptions, tasks, resources), and service categorization already exist, though in a very scattered format, in various service contract documents, common industry terminologies, and standardization templates such as ITIL, and COBIT. Knowledge-based approaches can utilize such information documented by domain experts. Consequently, we adopt a knowledge-based approach in developing the framework for service interdependency analysis.

Ontology-Based Information Extraction (OBIE)

In particular, we adopt Ontology-Based Information Extraction (OBIE) approach (Bontcheva 2004; Maynard 2005; Maynard et al. 2006). Ontological approaches are increasingly being used as formal knowledge representation techniques in extracting information from unstructured and semi structured domains due to their maintainability and reusability (Adrian 2008), and the advances in Semantic Web research (Davies et al. 2003; Davies et al. 2006; Fensel et al. 2003). OBIE has distinct advantages over traditional knowledge-based IE approaches (Maynard et al. 2005). In conventional IE approaches, domain concepts are captured and codified in gazetteers or lists. In case of OBIE, text occurrences of domain concepts are linked directly to their semantic descriptions in domain ontology. This allows tracing concepts across documents as well as reasoning with ontological relations and properties. From a technical standpoint, OBIE involves two main issues: (1) identification of concept instances from the ontology in the text, and (2) automatically populating ontology with concept instances in the text (Cunningham et al. 2006). OBIE approach has been used in market monitoring (Maynard et al. 2005) and business intelligence applications (Saggion et al. 2007).

Research Methodology

Given that the focus of our research is the design of an IT artifact, namely SLA-Miner; *Design Science* research methodology has been adopted. Particularly, we follow the seven guidelines for design science in Information Systems research prescribed by Hevner et al. (2004).

The developed prototype system, SLA-Miner, is the *instantiation of a product* (Gregor et al. 2007) and serves as a decision support tool for customer organizations using IT services (*Guideline 1: Design as an Artifact*). Literature in the practitioner community points toward the fact that interdependencies among IT service processes and their subsequent effect on KPIs often lead to sub-optimal outcomes and complications in both service delivery and performance monitoring (Hayes 2005; KPMG 2007; Munk 2006). A systematic way of addressing this issue by capturing and modeling such interrelationships is currently lacking (*Guideline 2: Problem Relevance*). The utility of SLA-Miner has been evaluated through use case scenarios and initial results are presented in this article (*Guideline 3: Design Evaluation)*. The SLA-Miner decision support system is unique in the sense that it focuses on knowledge extraction from SLA documents *prior* to service delivery, as compared to currently available tools that either lack in eliciting any such process knowledge or provide information *ex-post*, i.e., after service delivery *(Guideline 4: Research Contribution: Artifact)*. Additionally, the IT service ontology developed is a significant contribution of this research work, and encapsulates codified and semi-explicit knowledge from various resources to add to the foundation of design science knowledge base (Wand et al. 1993; Wand et al. 1995; Weber 1997) *(Guideline 4: Research Contribution: Foundations)*. The architecture of SLA-Miner is based on the principles in the area of Information Extraction, a rich area in artificial intelligence that is currently being applied in business intelligence generation (Cunningham et al. 2006; Hobbs et al. 2010; Maynard et al. 2005; Oleneme 2009; Saggion et al. 2007) (*Guideline 5: Research Rigor)*. Concepts and tools in the area of process management, IT service management, and information extraction were reviewed and evaluated, which finally resulted in the selection of Ontology-Based Information Extraction as the preferred approach for achieving the goal of eliciting process related information from text documents (*Guideline 6: Design as a Search Process)*. Finally, our research provides tangible results for the practitioner community, while also contributing to the research efforts toward SLA automation, and knowledge representation techniques, through the SLA-Miner decision support tool and the associated IT service ontology development (Paschke et al. 2008) (*Guideline 7: Communication of Research).*

In the following section, we illustrate how these design science principles have helped us in developing the architecture for SLA-Miner based on theories of Ontology-Based Information Extraction.

Architecture of SLA-Miner

Figure 1 shows the architecture of SLA-Miner. As depicted in the figure, SLA-Miner consists of following main components: (1) *SLA Entity Recognition*, (2) *SLA Context Inference*, and (3) *SLA Analytics*. Each of these components is detailed next.

SLA-Miner Component 1 – SLA Entity Recognition

Component 1 is concerned with the task of identifying domain entities from each SLA in the corpus of IT service contract documents. This involves extracting key entities that characterize the service (such as service name, and performance measures), as well as provide information from a business process perspective (such as activities, events, and resources). It also involves finding aliases (such as *Networking Service* also referred to as *Internet Access Service*), as well as finding basic context information for each SLA entity (such as the service name and performance measure name in whose context each service entity is mentioned). This refers to NE, CO, TE, and TR (partial) tasks, from an IE standpoint (MUC-7 1998). The goal of identifying SLA entities and basic context information is to provide a starting point to extract additional context information and draw inferences about these entities in the following stages.

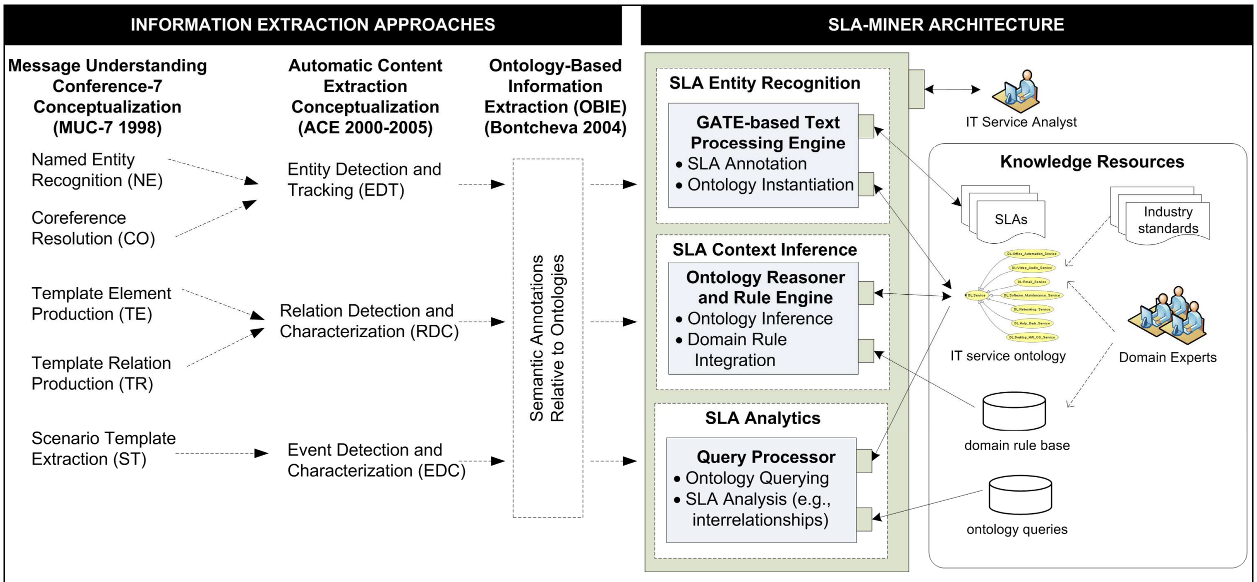

Figure 1: SLA-Miner Architecture

First, a unified IT service ontology that captures service-related concepts and their associated hierarchy, relations, and properties has been developed. The IT service ontology represents the IT service domain knowledge that helps the identification of service entities from SLA text documents. At the core of the IT service ontology are classes representing 1) services (e.g. *Office Automation Services*, *Help Desk Service*, *Network Service*) from the service description, 2) service performance measures/KPI associated with each service, and 3) activities, events and resources associated with each service performance measure (e.g. *Configuration* activity, *Installation* activity, *Trouble Ticket* event) (Green 2006; Hill 2006). This ontology creation has been guided by terminologies used in IT service management and business objectives (Buco et al. 2004; Hasselmeyer et al. 2007). In constructing the service taxonomy, we used a variety of practitioner resources - especially industry standards such as ITIL (2010) and Open Group (2010). We also utilized the SLA Management Handbook (2004), real-world SLA documents accessible to us, service classifications used at different academic institutions (available through a colleague at the ITIL academic forum), and SLAs available on the web. The aim is twofold. First, we use industry standard terminology so that the terms are recognizable to SLA-Miner. Second, we utilize the semi-formal knowledge structure that already exists in the IT service management domain so that the ontology is closely aligned to the industry practice.

Next, for annotating the SLA documents, SLA-Miner utilizes an open source toolkit for basic language processing, GATE (General Architecture for Text Engineering), developed at the University of Sheffield, which includes a skeleton IE system called ANNIE (A Nearly-New Information Extraction System) (Cunningham et al. 2002). It also supports a knowledge-based IE approach and provides application programming interfaces (APIs), which were used in the development of SLA-Miner.

The SLA documents and the IT service ontology serve as the language resources (inputs) to this SLA-Miner component. Commonly used IE processing modules supported by GATE, including tokenizer, sentence splitter, part of speech tagger, stemmer were used. Additionally, to enable OBIE, we used ontology-based gazetteers, Apolda (Automated Processing of Ontologies with Lexical Denotations for Annotation) plugin for GATE (Wartena et al. 2007), and JAPE (Java Annotations Patterns Engine) grammar rule language (Cunningham et al. 2000). The ontology-based gazetteers is better suited for extracting fewer number of concepts (such as service names and performance measure names), while Apolda is better suited for extracting large number of concepts with only a few textual representations available per concept (such as different activities, events, resources). Together, these two resources help accomplish NE, CO, and TE tasks, as mentioned in the MUC-7 (1998) conceptualization.

The SLA Entity Recognition component also extracts basic context information for each of the annotated service entity. This information includes the context in which each service concept is mentioned in a SLA document, such as the containing service and the performance measure names. Note that the SLA documents are input in XML format with tags indicating text segments such as service name, service description, and KPI. JAPE grammar rules (based on pattern-matching) are developed for analyzing the annotation information (start and end nodes) derived from these XML tags in association with the extracted service entities to derive and record the context information.

The resultant annotations for the service entities in the SLA documents consist of feature information such as the name of the matched ontology class, the ontology URI, start node of annotation, end node of annotation, annotated phrase, service name in which the phrase is mentioned, and KPI name in which the phrase is mentioned. Upon annotating the SLA documents with entity and basic context information, another set of JAPE rules are used to populate the IT service ontology by transferring the annotations from SLA documents in the form of ontology individuals. Thus, after ontology population, each individual in the ontology represents an annotation of an ontology class. The annotation features are transferred in the form of datatype properties (such as *inContextOfService*, and *inContextOfServiceQuality)* of the IT service ontology classes for subsequent semantic processing.

SLA-Miner Component 2 – SLA Context Inference

Component 2 is concerned with the task of extracting additional context information for the service entities, and accomplishing the TR task mentioned in the MUC-7 (1998) conceptualization. First, the inferred class and individuals hierarchy is computed for the populated IT service ontology using an ontology reasoning engine. SLA-Miner utilized an open source reasoned called Pellet for this purpose. Pellet performs subsumption reasoning, and the resultant inferred class and individual hierarchy forms the knowledge base for further semantic inferencing in performing OBIE.

Next, a set of domain rules are developed specifying the relationships between service entities that are typically observed. For example, *Installation Accuracy* is a performance measure (KPI) used for an *Office Automation Software Service*. This kind of domain knowledge is elicited from experts and encoded as logical rules in the Semantic Web Rule Language (SWRL). Each rule represents an if-then clause, and new knowledge is added to the knowledge base if the rule condition is matched. For example, if *Installation* activity is observed to be mentioned in the context of *Installation Accuracy* KPI, then for those individuals representing the *Installation Accuracy* KPI, an object property (*mentionsAssociatedActivity*) is populated indicating that an ‘associated’ activity is observed. Similarly, rules can also help identify whether any unrelated activities are mentioned in the context of a KPI, and populate another object property (*mentionsOtherActivity*). This object property population technique based on domain rules is used to extract information about relationships between Service-KPI, KPI-KPI, KPI-Activity, KPI-Resource, and KPI-Event pairs. SLA-Miner uses Jess rule engine for drawing inferences based on rules.

SLA-Miner Component 3 – SLA Analytics

Component 3 is concerned with identifying analytic information, across service contract documents (such as *Office Automation Software Service* relies on *Help Desk Service* in managing *Trouble Ticket* events). This refers to a complex IE task, namely ST (MUC-7 1998), and is domain-dependent. Intelligence information is derived by constructing scenarios of interest or by making inferences based on basic analytic relations among entities identified in the preceding stages (Cunningham et al. 2006).

A set of interesting analytical queries are developed, encoded in the form of Semantic Query-enhanced Rule Language (SQWRL), and stored in a query base. SQWRL provides advanced querying capabilities for retrieving knowledge from ontologies through a SQL-like query language. In SQWRL, the SWRL rule antecedents are used as pattern specifications, while the SWRL rule consequents are replaced with selection, collection, and formatting operators. In SLA-Miner, the Jess rule-based engine is used to execute SQWRL queries, similar to SWRL rules.

Next, SLA analysis involves consuming the query results by the Service Analyst for discerning the interrelationships between different IT service entities and using the analytic information for applications such as contract negotiations, incentive and process management decisions during the IT service delivery and management process. Next section, which discusses preliminary validation through use cases, provides examples of analytic information derived through the use of SLA-Miner.

Preliminary Validation

The current version of SLA-Miner utilizes 7 SLAs selected from the set of 37 SLAs (mentioned earlier in Section 1) for pilot testing the prototype. We generate illustrative scenarios representative of the tasks normally undertaken by a manager at the customer organization responsible for coordinating the contract. Such scenario generation is a well accepted method of validation in design science research (Paschke et al. 2008; Wang et al. 2009). Some illustrative scenarios are presented below.

- *How are different services related to one another? Which activities/resources/events have the most impact on different services/i.e. which activities/resources/events impact more than one service and to what extent?* Identification of such causal relationships not only helps in better process design, but also in root-cause analysis in case of repeated failures.
- *Which KPIs are most important? How does the causal relationships affect (multiple) performance measure(s)?* This knowledge is very important in formulating the right incentive-based compensation structure. For example, if higher incentives are placed on the speed of problem resolution than on resolving/preventing the root cause; the provider organization may not make its staffing and effort allocation decisions in the best interest of the customer.

Panels 1, 2, and 3 of Figure 2 provide a sample of the aggregate information extracted from the source SLA document[1]. Panel 1 displays the interrelationship among the different services at a high level (*even though it is not shown in the diagrams, each interrelationship can be tracked down to the activity/resource/event that is causing it*). It is apparent from the figure that while each service is related to every other service, *Desktop Video Teleconference (VT)* has the highest linkage with others in the bundle, followed by *E-mail Service* and *Office Automation Software Service*. On the other hand, *Help Desk* and *Internet Access* services have relatively less impact on others. The fact that *Desktop VT* has the highest linkage with others was surprising at first. However, closer investigation (not shown in Figure 2) reveals that this is due to the fact that this service shares many resources (*Hardware*, *Software*, *Data*, and *Network*) and an activity (*Establish Connection*) with others.

Panel 2 of Figure 2 shows the relative importance of the KPIs in terms of both their linkage with and impact on other services. *Interoperability* emerges to be the most important KPI as it has the highest impact and is also related to most other services. The KPI *Availability*, while also linked to most other services has lesser impact. On the other hand, KPIs like *Installation Accuracy* and *Upgrade Currency*, while not liked to more than one/two services, has a higher impact on the services they measure. The implication for such interrelationship is that the *Interoperability* dimension should be measured and monitored more closely and higher incentives should be provided for this performance measure. On the other hand, some sort of balance must be drawn while deciding on the incentive level between *Availability* and *Installation Accuracy*/*Upgrade Currency*.

Finally, Panel 3 of Figure 2 illustrates how the different activities affect different IT Services. From the analysis, *Survey* and *Installation* emerged. Interestingly, activities like *Monitor*, *Restore*, and *Respond* seem to have very little linkage with other services. While this may be idiosyncratic to the current SLA we used, it serves as an important checkpoint as to whether these activities, especially *Monitor*, are sufficiently stressed in the document, and whether any modifications are needed.

---

[1] Using the SLA-Miner, we were able to extract information at the smallest level of granularity; only aggregate information is shown due to clarity and space concerns

Conclusion and Future Work

In this paper, we propose an Ontology- Based Information Extraction approach for identifying KPI interrelations in a multi-service IT outsourcing contract from textual Service Level Agreements. The purpose of our research is to assist the participants of an IT service arrangement in obtaining a better appreciation of the interrelationships among different service components so that measurement, monitoring, and subsequent compensation schemes are more effective for both organizations. To this end, we have developed and presented a prototype in the current paper – SLA-Miner – as an instantiation of the framework that utilizes a set of real-life SLAs.

The contribution of our work is two-fold. To the best of our knowledge, this is the first effort toward systematically identifying and categorizing service process interrelationships from SLA documents. This has tremendous implication for customers of IT service processes as it provides a toolset to uncover and understand hidden relationship patterns. Second, the service ontology and associated rule base provides structure and codification to the IT Service domain knowledge, which can be utilized in frameworks like ITIL and COBIT.

Our research will contribute to the recent body of knowledge in the area of business intelligence generation from process mining (e.g. Li et al. (2009)) and automated SLA Management (e.g., Paschke et al. (2008)). Our ongoing and future research agenda involves validation of the information extracted using multiple domain experts, as well as ensuring the robustness of the IT service ontology and rule base using SLAs from different organizations and industry sectors.

Figure 2: Preliminary Results for Illustrative Scenarios

References

ACE "Automatic content extraction," http://www.ldc.upenn.edu/projects/ACE, 2000-2005.

Adrian, B. "Preface - Proceedings of Ontology Based Information Extraction Systems (OBIES 2008)," 2008.

Appelt, D. "An introduction to information extraction," *Artificial Intelligence Communications* (12:3) 1999, pp 161-172.

Bontcheva, K. "Open-source tools for creation, maintenance, and storage of lexical resources for language generation from ontologies," Proceedings of 4th Language Resources and Evaluation Conference (LREC'04). 2004.

Buco, M.J., Chang, R.N., Luan, L.Z., Ward, C., Wolf, J.L., and Yu, P.S. "Utility computing SLA management based upon business objectives," *IBM Systems Journal* (43:1) 2004, pp 159-178.

Chen-Yu, L., and Von-Wun, S. "Ontology-based information retrieval and extraction," Information Technology: Research and Education, 2005. ITRE 2005. 3rd International Conference on, 2005, pp. 265-269.

Cleary, R. "IT Measures That Matter," in: *Information Week*, 2008, pp. 53-56.

Cowie, J., and Lehnert, W. "Information extraction," *Communications of the ACM* (39:1) 1996, pp 80-91.

Cunningham, H., and Keith, B. "Information Extraction, Automatic," in: *Encyclopedia of Language & Linguistics*, Elsevier, Oxford, 2006, pp. 665-677.

Cunningham, H., Maynard, D., Bontcheva, K., and Tablan, V. "GATE: A framework and graphical development environment for robust NLP tools and applications," Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL '02), Philadelphia, PA, 2002.

Cunningham, H., Maynard, D., and Tablan, V. "JAPE: A Java Annotation Patterns Engine," CS-00-10, Department of Computer Science, University of Sheffield.

Davies, J., Fensel, D., and van Harmelen, F. (eds.) *Towards the Semantic Web: Ontology-driven Knowledge Management*. John Wiley and Sons, 2003.

Davies, J., Studer, R., and Warren, P. (eds.) *Semantic Web Technologies: Trends and Research in Ontology-based Systems*. John Wiley and Sons, 2006.

Edgington, T.M., Raghu, T.S., and Vinze, A.S. "Using process mining to identify coordination patterns in IT service management," *Decision Support Systems* (In Press, Corrected Proof) 2010.

Fensel, D., Hendler, J., Lieberman, H., and Wahlster, W. (eds.) *Spinning the semantic web: Bringing the world wide web to its full potential*. The MIT Press, Cambridge, MA, 2003.

Fitoussi, D., and Gurbaxani, V. "IT Outsourcing Contracts and Performance Measurement," Uiniversity of California Irvine: The Center for Research in Information Technology and Organizations, 2008.

Green, L. "Service level agreements: an ontological approach," 8th international conference on Electronic commerce: The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet, ACM, 2006, pp. 185 - 194

Gregor, S., and Jones, D. "The anatomy of a design theory," *Journal of the Association for Information Systems* (8:5) 2007, pp 312–335.

Grigori, D., Casati, F., Castellanos, M., Dayal, U., Sayal, M., and Shan, M.C. "Business process intelligence," *Computers in Industry* (53:3), Apr 2004, pp 321-343.

Hasselmeyer, P., Koller, B., Kotsiopoulos, I., Kuo, D., and Parkin, M. "Negotiating SLAs with dynamic pricing policies," in: *Proceedings of the Service Oriented Computing: a look at the Inside (SOC@Inside '07)*, 2007.

Hayes, I.S. "Metrics for IT Outsourcing Service Level Agreements," Clarity Consulting.

Hevner, A.R., March, S.T., Park, J., and Ram, S. "Design science in information systems research," *MIS Quarterly* (28:1) 2004, pp 75-105.

Hill, M. "Service Taxonomy and Service Ontologies Deliver Success to Enterprise SOA," in: *SOA World Magazine*, 2006.

Hobbs, J.R., and Riloff, E. "Information extraction," in: *Handbook of Natural Language Processing,* N. Indurkhya and F.J. Damerau (eds.), CRC Press, 2010.

ITIL "http://www.itil-officialsite.com/home/home.asp," 2010.

KPMG "Strategic Evolution A Global Survey of Sourcing Today " in: *KPMG: IT Advisory*, 2007.

Li, J., Wang, H.J., Zhang, Z., and Zhao, J.L. "Mining business policy texts for discovering process models: A framework and some initial results," in: *Proceedings of the Sixth Workshop on e-Business (WeB 2007)*, Montreal, Quebec, Canada, 2007a.

Li, J., Wang, H.J., Zhang, Z., and Zhao, J.L. "A policy-based process mining framework: mining business policy texts for discovering process models," *Information Systems and E-Business Management*) 2009.

Li, Y., and Bontcheva, K. "Hierarchical, Perceptronlike Learning for Ontology-Based Information Extraction," 16th International World Wide Web Conference, Banff, Canada, 2007b, pp. 777-786.

Mani, D., Barua, A., and Whinston, A. "Successfully Governing Business Process Outsourcing Relationships," *MIS Quarterly Executive* (5:1) 2006, pp 15-29.

Maynard, D. "Benchmarking ontology-based annotation tools for the Semantic Web," in: *Proceedings of UK e-Science Programme All Hands Meeting (AHM2005) Workshop: Text Mining, e-Research and Grid-enabled Language Technology*, Nottingham, UK, 2005.

Maynard, D., Peters, W., and Li, Y. "Metrics for evaluation of ontology-based information extraction," in: *Proceedings of WWW 2006 Workshop on Evaluation of Ontologies for the Web (EON 2006)*, Edinburgh, Scotland, 2006.

Maynard, D., Yankova, M., Kourakis, A., and Kokossis, A. "Ontology-based information extraction for market monitoring and technology watch," in: *Proceedings of the ESWC Workshop: End User Aspects of the Semantic Web*, Heraklion, Crete, 2005.

MUC-7 *Proceedings of the 7th Message Understanding Conference (MUC-7)* Morgan Kaufmann, San Mateo, CA, 1998.

Munk, C.W. "More Companies Offer Incentives To Outsourcers," in: *Wall Street Journal*, 2006, p. 1.

Oleneme, D.U. "Information Extraction: Extraction of entities, relations, events and facts from bankruptcy newswire corpora," in: *School of Computer Science*, The University of Manchester, 2009.

Open Group "SLA Management Handbook: Enterprise Perspective, Volume 4," The Open Group, p. 137.

OpenGroup "http://www.opengroup.org/," 2010.

Paschke, A., and Bichler, M. "Knowledge representation concepts for automated SLA management," *Decision Support Systems* (46:1) 2008, pp 187-205.

Pazienza, M.T. (ed.) *Information extraction in the Web era*. Springer-Verlag, Berlin, 2003.

Saggion, H., Funk, A., and Maynard, D. "Ontology-based Information Extraction for Business Intelligence," in: *6th International Semantic Web Conference (ISWC 2007)*, Busan, Korea, 2007.

Taylor, R., and Tofts, C. "Death by a thousand SLAs: A Short Study of Commercial Suicide Pacts," HP.

van der Aalst, W.M.P. "Business alignment: Using process mining as a tool for delta analysis and conformance testing," *Requirements Engineering* (10:3), Nov 2005, pp 198-211.

Wand, Y., and Weber, R. "On the Ontological Expressiveness of Information Systems Analysis and Design Grammars," *Information Systems Journal* (3:4) 1993, pp 217-237.

Wand, Y., and Weber, R. "On the deep structure of information systems," Blackwell Publishing Ltd, 1995, pp. 203-223.

Wang, H.J., Zhao, J.L., and Zhang, L.-J. "Policy-Driven Process Mapping (PDPM): Discovering process models from business policies," *Decision Support Systems* (48:1) 2009, pp 267-281.

Wartena, C., Brussee, R., Gazendam, L., and Huijsen, W.-O. "Apolda: A practical tool for semantic annotation," Proceedings of the 18th International Conference on Database and Expert Systems Applications (DEXA '07), 2007, pp. 288-292.

Weber, R. *Ontological Foundations of Information Systems*, 1997.