# KDD and Its Applications in Automotive Sector - A Brief Survey.

**Conference Paper** · January 2008

Source: DBLP



Setu Namburu
Northwestern University



Liu Qiao
Beijing University of Chemical Technology



Krishna R. Pattipati
University of Connecticut

# KDD and Its Applications in Automotive Sector – A Brief Survey

**Setu Madhavi Namburu[1], Danil Prokhorov[1], Shunsuke Chigusa[1], Liu Qiao[1], and Krishna Pattipati[2]**
[1]Toyota Motor Engg. & Mfg. North America, Ann Arbor, MI
[2]Dept. of ECE, Univ. of Connecticut, Storrs, CT

***Abstract -*** *Modern information society is facing many challenges due to the exploding sources of heterogeneous information. The typical of them include managing the data, integrating disparate data sources, imputing missing values in the data, extracting meaningful information from structured/unstructured data, developing tools for data analysis and predictive models, to name a few. The information technologies, such as data-mining, text-mining and semantic web, are playing major roles in tackling these issues. Automotive industry is one of the examples, where these technologies are not as deeply explored as other areas, but they could play a vital role in managing/analyzing automotive data. In this paper, we provide a brief survey of KDD technologies and their application to automotive sector, with particular attention to automotive service.*

***Keywords:*** *Engineering Data, Automotive Service, Data Mining, Text Mining, Data Management, Semantic Web.*

## 1 Introduction

In any enterprise, irrespective of its size and type, the amount of data that one has to deal with is getting larger every year due to widely computerized operations. The data could include any/all of numeric, text, metadata (XML, RDF, etc) and multi-media (audio, video, images). The Knowledge Discovery and Data-mining (KDD) is one of the fastest growing fields in this era of "Information Age". Several theories and software tools have emerged and are playing a key role in dealing with growing sources of information/data. Fayyad, *et al* [23] defined KDD as a "nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data".

Industries, such as automotive and aerospace, deal with terabytes of product-related data from design to after-sales. There is a crucial need for intelligent tools to integrate data from disparate sources (across the organization, e.g., data from design, manufacturing, operations, training and service), manage, impute/process/analyze data (multimedia / numeric / text / metadata), bring structure to it, and discover knowledge from this data. In an effort to investigate the technologies available in the KDD community and the benefits it can provide to automotive industry, this paper provides a brief literature survey of KDD technologies as applied to automotive sector, with particular attention focused on automotive service. In recent years, integration of several key technologies (data-mining, text-mining, knowledge discovery, etc.) has become important to address real-world challenges successfully.

Here, we convey the survey from three view points, considering all of them under the rubric of KDD.

1. Modeling and management of the engineering data repositories;

2. Application of data mining tools in the automotive sector; and

3. Application of text mining tools/text analytics to the automotive sector.

The rest of the paper is organized as follows. Section 2 provides a short description of KDD technologies and their applications, in general. In section 3, we give a brief introduction to automotive service sector and data challenges one has to deal with. Next, we provide a survey of KDD applications in the automotive domain in section 4. Finally, in section 5, we conclude the paper with a brief discussion of practical opportunities & challenges for the KDD research community in the automotive sector.

## 2 KDD Technologies

Enormous amount of heterogeneous data is the backbone for technological innovations in the KDD community. KDD can be considered as an interdisciplinary field at the intersection of several related areas: information retrieval, artificial intelligence, machine learning, statistics, natural language processing, and data mining. It is often difficult to define the terms, data mining, text mining, and knowledge discovery individually in this multi-facet information era. There appeared an article on Duo-mining (combining data and text-mining), describing the benefits and necessity of combined technology [24]. Almost all kinds of businesses benefit from the KDD technologies, ranging from insurance, finance, information technology, design, manufacturing, after-market, business intelligence, bioinformatics, medicine, marketing and so on. The example data-sources range from simple text documents to complex data sources (consisting of multi-media data).

A useful data mining process model that is industry-neutral and tool-neutral was proposed by CRISP-DM (Cross-Industry Standard Process for Data Mining) group (Figure 1). Their design was based on a study supported by a consortium of automotive, aerospace, telecommunications and consultancy companies, including DaimlerChrysler, OHRA (insurance company), NCR (data warehouse supplier), and integral solutions limited [22]. Cios and Kurgan provided a detailed description of a six-step process of the CRISP-DM model, based on technologies such as XML, PMML, SOAP, UDDI, and OLE BD-DM and their implementation details [21]. Figure 2 shows the evolution of data-mining theories and applications [21].

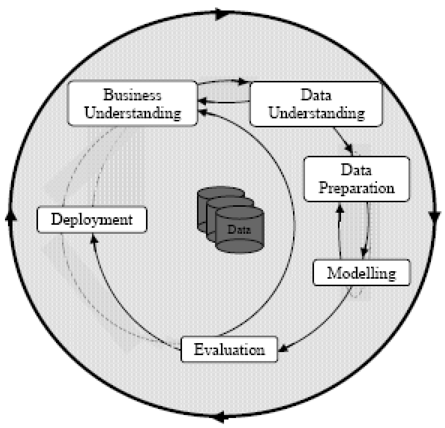


Figure 1. CRISP-DM process model [22]

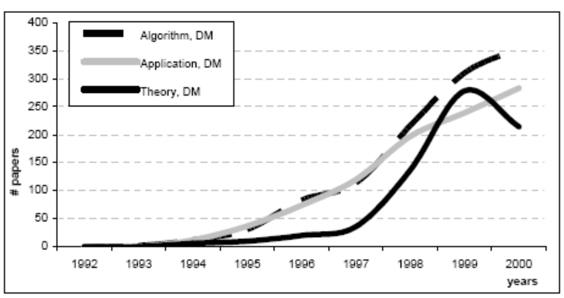


Figure 2. Evolution of data mining theory and applications [21]

The most important step in dealing with data is standardizing its structure and knowledge management. Using the so-called metadata provides an opportunity to reuse existing information frequently. XML has been widely used and accepted by the computer industry, to represent not only human-readable documents, but data in general. The XML standards give a syntactic structure for describing data. Many technologies, such as SGML, HTML, XML, RDF, and OWL, have evolved with the advent of WWW. Over time, semantic web has become a popular choice in enterprises, due to its ontological description of knowledge. Ontology is "a formal explicit specification of a shared conceptualization" and is analogous to relational database. As stated in Wikipedia, "Semantic web is an evolving extension of the WWW in which meaning of information and services on the web are defined, making it possible for the web to understand and satisfy the requests of people and machines to use the web content". Figure 3 shows a famous "semantic web layer cake". Recently, the WWW consortium (W3C) has published two significant semantic web specifications: the first is the query language specification, SPARQL, and second the Ontology Language for the Web (OWL) 1.1 [26].

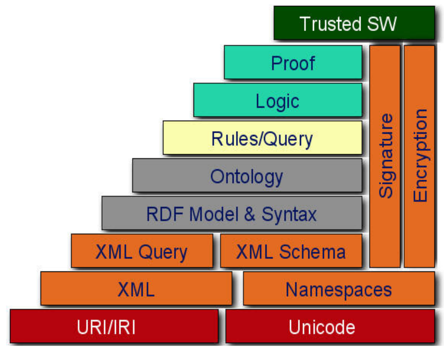


Figure 3. Semantic web layer cake [26]

Next comes the analysis/processing of data. Goebel and Gruenwald provided a survey of a wide range of KDD software tools and their capabilities and a brief review of existing data-mining methods [17]. Hotho *et al* provided an excellent survey on text-mining, considering it as an interdisciplinary method encompassing information retrieval, machine learning, statistics, computational linguistics and data-mining [2]. It is estimated that 85% of business information lives in the form of text (mostly natural and unstructured text). The readers are referred to the literature on data-mining/text-mining methodologies, and their wide range of applications.

# 3 Automotive Service Sector

The factors, such as stricter environmental regulations from the government (e.g., exhaust emissions, fuel consumption), legal regulations for safety and customer demands for enhanced features are the main reasons for the rapid growth in vehicle complexity. To meet the market demands on time, the new designs have to be tested and implemented as rapidly as possible, resulting in the shortening of automotive development cycles. These factors, in turn, are forcing the automotive manufacturers to rely on reusable and standardized tools (e.g., reusable software, development of standardized structures for data exchange, standardized design tools, etc). One of the primary problems faced by vehicle engineering design teams is one of improving the use of available engineering data and procedural knowledge, whose management (modeling, integration, sharing and re-use), adds a well recognized value [12]. The data sources in automotive industry include, engineering data (CAD data, metadata, office documents, etc.), field data in several forms (sensor data, data in text that contains valuable diagnostic information, such as case bases & warranty data, marketing information, consumer evaluation, vehicle satisfaction details, etc).

With the growing complexity of vehicles, diagnostics is becoming increasingly important and considerably complex. The pressure on automotive service technicians is increasing to quickly and accurately repair vehicles, and, at the same time, keep up with the ever changing developments in vehicle technologies. Hilger et al presented major categories of concerns and diagnostic challenges in an automotive workshop [29] (Figure 4). The globalization of product development is only leaving complicated challenges to manufacturers, even though considerable cost savings are achieved.


Figure 4. Major categories of concern to dealer technician [29]

The patent from Venkataraman *et al* [19] describes the knowledge stores, interactive diagnostic services, and interactive diagnostic systems. Interactive answers to the questions are used to dynamically mine the data store for potential causes and for corrective actions to the problem. Figure 5 show the typical data sources involved in building the knowledge store [19].

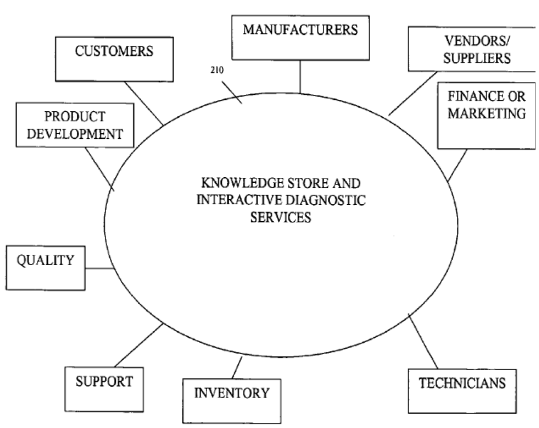


Figure 5. The knowledge store for interactive diagnostic services [19]

The data-mining/text-mining technologies such as data-pre-processing, classification, regression, case based reasoning, data visualization, predictive analytics, natural language processing bring several advantages to the automotive service sector. Even though industries, such as aviation/aerospace are far ahead in adapting several standardized languages (FARs, NOTAMs, ADs, AECMA, ATA-100 etc), automotive industry is still lagging behind considerably in using the advanced KDD technologies. Engineering and information technology group at Boeing is involved in developing technologies for information management and collaborative technologies. DARPA's CBM+ (condition based maintenance plus) initiative opened many avenues for applied research based on data-mining/text-mining technologies. NASA is conducting significant research in health management technologies (KDD technologies are widely used for remote diagnostics and prognostics). As remote diagnostics is gaining momentum in terms of customer reception, many automotive companies are investing considerable research efforts in mining the field data. We believe that it is time for the automotive industry to take part in the KDD activities to achieve considerable cost and time savings and improved customer satisfaction.

# 4 Survey of KDD Applications in Automotive Sector

This section provides a brief literature survey of data management, data-mining/text-mining technologies and their applications to real-world automotive problems. Significant amount of literature has been published by research organizations and automotive companies, describing their applications. The purpose here is not to provide an exhaustive literature review, but rather to briefly describe relevant case studies to indicate how automotive sector can benefit from the KDD technologies.

One of the key aspects is the standardization of service information in the automotive industry. In an effort to achieve standardization, a task force has been created by the SAE (Society of Automotive Engineers) in the United States. The result was the J2008 standard, which was released in 1997. A logical relational model was created for data modeling and SGML was used as the standard for exchanging the service information. In the same vein, with the co-operation of many automotive parties in Europe, a specification called OASIS [25] was developed to provide a standard format for the service information. The technical framework for representing the metadata was specified based on existing recommendations from the World Wide Web Consortium (W3C) - the RDF [26].

Until now, in the automotive field, solutions for modeling the data involved RDBMS (Relational Data Base Management Systems). The requirements for its object model specification and predefining the data structures prevent the flexibility required for management of data, such as automotive engineering/service data. Semantic web technologies (RDF, OWL, etc) promise to provide advanced capabilities for data management.

In 2006, Renault built a probabilistic induction tool aimed at diagnostics [11, 28] (Figure 6). The goal was to show that, rather than having all of the procedures written in repair manuals as they are today, a web engine could compute them on the fly, to minimize the cost of diagnostics. OWL was used to model the repository and the related concepts, and RDF for information exchange [11].

At Elasis S.C.P.A (A Fiat group company), semantic web technologies have been tried for testing in the automotive product development. Macrini *et al* [12] demonstrated the case study of road tests in automotive industry, by building the semantic information system using RDF. The solution was found in:

A. The design of a specific ontology for test data;

B. An agent called "spider" searches autonomously what is of interest from a list of locations in the company intranet, analyzes its contents and defines RDF/XML statements according to ontology;.

C. A web browser to navigate the available set of RDF/XML statements; and

D. A neural network designed to classify data to identify and suggest analogies between test results.


Figure 6. Screenshot of the system. Two successive commented screenshots (1st Phase and 2nd phase) show the system's usage in more details [28]

The semantic web technology has been tried not only for service data management, but also for improving the quality and efficiency of innovative product design – the WIDE project [27] (See Figure 7). WIDE aims to improve the product design by applying emerging semantic web (SW) techniques to develop and test an effective information management and knowledge sharing system for multi-disciplinary design teams.

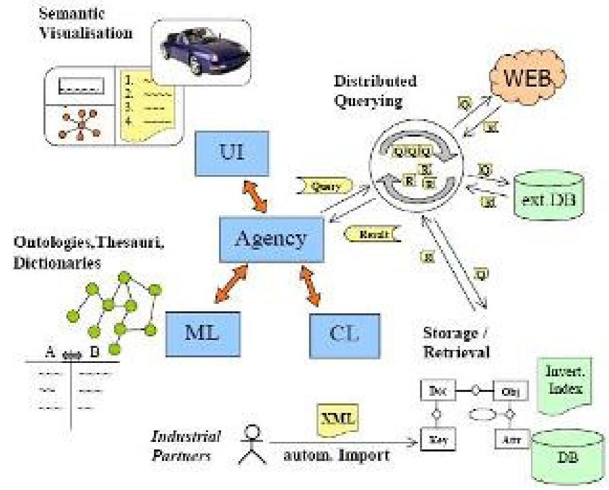


Figure 7. Diagram showing the WIDE Architecture [27]

Ahmed *et al* [8] proposed a methodology for the development of an integrated taxonomy and demonstrated through a case study of "development of ontology for the purpose of indexing, searching and retrieving engineering design knowledge" (EDIT – Engineering Design Integrated Taxonomy). This research brought together knowledge from researchers undertaking empirical research in engineering design, taxonomies from literature and computational techniques particularly in the field of natural language processing tools. Stock *et al* [10] from the BMW group presented the concept for a generic document model (for repair technical documentation), and an authoring environment based on the separation of content from representation. To achieve this goal, the instructions were visualized by technologies, adding dynamic information as further dimension to the heretofore static information. They proposed to use Augmented Reality (based on ARVIKA project [16]) to support service personnel in carrying out their tasks by displaying context-sensitive, additional and virtual information in their field of view, e.g. via see-through data goggles (Figure 8). In addition to Augmented Reality, other potential technologies are videos, animations, hypermedia and virtual reality (potential multi-media applications).

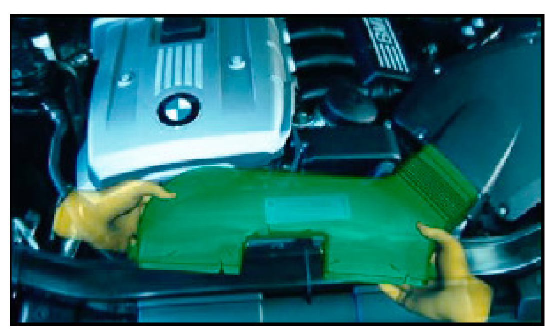

Figure 8. Augmented reality scenario [10]

Next, data-mining/text-mining technologies provide numerous advantages in analyzing the service data. It is estimated that world's carmakers spend 25-30 billion dollars on warranty claims each year. This shows an urgent need for effective data mining tools. Many software tools (SAS enterprise miner, SPSS, ThemeScape/ThemeRiver, PolyAnalyst, etc.) have emerged over the years and are offering a large scale data-mining/text-mining solutions with provision of easy-to-use interfaces. Companies, such as Volvo and Honda, are widely using SAS Enterprise-miner & Text-miner to analyze warranty data [5,6,7]. Ville [6,7] presented warranty code predictive classification using SAS Text-miner by extracting the information from the text fields of over 1 million records. This work demonstrated that thousands of warranty codes could be reliably assigned by computer agents with an accuracy rate that ranged from 60% to 90% (compared to human accuracy rates in the 50% range). Wallace *et al* [5] used SAS text miner on warranty and call center data for early warning for product quality awareness. Many research organizations are actively producing free software (OntoGen, Stanford Parser, Protégé, WordStat/SimStat, RapidMiner, Bow Toolkit, GATE, SOM Toolbox, etc.) for several applications, such as data/text preprocessing, visualization, ontology generation, classification, regression tools compared to expensive commercial software tools. OntoGen is one of the examples of free software tools produced by researchers from Jozef Stefan institute [18]. The main features of the system include unsupervised and supervised methods for concept suggestion and concept naming, as well as ontology and concept visualization. Figure 9 gives a glimpse of the OntoGen system. Considering the unstructured nature of information from disparate sources, this kind of tools are gaining momentum in producing great results for real-world academic and enterprise applications. Thomas Montgomery, a technical expert from the Ford motor company, says a budgeted way of doing text-mining is by following the "Reduce, Reuse, Recycle" principle [32]. He showed a text-mining case study on eSigma problem to quickly understand thousands of customer verbatims from warranty claims.

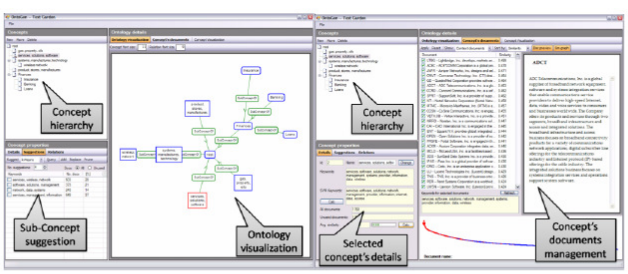


Figure 9. The OntoGen system for ontology generation and visualization [18]

The diagnostic data from the field typically consists of sensor data from the vehicle components, warranty data, technical assistance case bases (the notes are made in natural language while communicating with the technicians during diagnosis), and customer relations databases. The mining of these data sources by processing and discovering hidden knowledge from unstructured data sources could save significant dollars to the service sector. This line of research gave birth to dynamic case-based reasoning (CBR), where many researchers used technologies from artificial intelligence, information retrieval and natural language processing to solve real-world troubleshooting problems. CBR helps in troubleshooting a problem by comparing new cases with existing cases and producing decision rules in aiding the technicians. Saxena [13] proposed an integrated reasoning architecture for fault diagnosis and prognosis in industrial environments, confirming with the standard case-based reasoning cycle – comprising the *four Rs*: Retrieve, Reuse, Revise and Retain. According to this architecture, the diagnosis is carried out at two levels, first by using the textual observations (initial diagnosis for fault localization) and second by using relevant sensor data and analytical techniques (final diagnosis for fault detection and isolation) as shown in Figure 10.

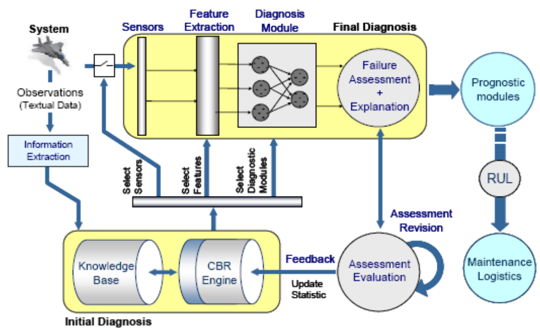


Figure 10. Integrated reasoning architecture for fault diagnosis and prognosis in industrial environments [13]

Huang and Murphy [1] presented an application of text-mining technology for automatic mapping of problem descriptions to the correct diagnostic categories using many term weighting schemes and compared the results with an LSA based method and human reasoning. Jarmulak et al presented methods for one important aspect, easing or automating the acquisition of adaptive knowledge for CBR design [9]. Gersten *et a*l [22] from DaimlerChrysler used a commercial data-mining tool, Clementine, for predictive modeling in automotive direct marketing (by considering data from acquisition campaigns). Keller *et al* applied the KDD methods (based on decision rule learner with multivariate discretization for classification and M6 for regression learning) for automotive data engineering mechanisms and processes in vehicle design and compared their performance with neural network based techniques [20]. Many other algorithms based on hidden markov models, decision trees, fuzzy rules, graphical models, support vector machines, etc. have also been used in the automotive sector to deal with the field data.

# 5 Conclusion

We would like to conclude the paper with a brief discussion of opportunities and challenges for the KDD community in applying the technologies to the automotive sector. The increasing complexity and growing electronic content of modern vehicle has made diagnostics a significant challenge to the automotive industry. It is estimated that in the past decade, there is more than 300% increase in the complexity of electronics in vehicles [29] (Figure 11). Modern vehicles are expected to contain more than 50 ECUs (Electronic Control Units). Conventional design of products used simple maps and rule-based systems for control and diagnostic design. Modern control/diagnostic systems use advanced techniques (analytical redundancy, machine learning, decision trees, statistical classification tools, fuzzy inference systems, etc.). With the advancements in vehicle technology, the vehicle diagnostics has become a puzzle for average technicians. With the advent of remote diagnostics (e.g., OnStar), data-mining and text-mining technologies have become necessary tools to deal with terabytes of field data on a daily basis. The demand for reusable software, standardization of knowledge management, advanced text analytics to deal with unstructured field data is increasing rapidly with the release of many types of vehicles into the market, every year.

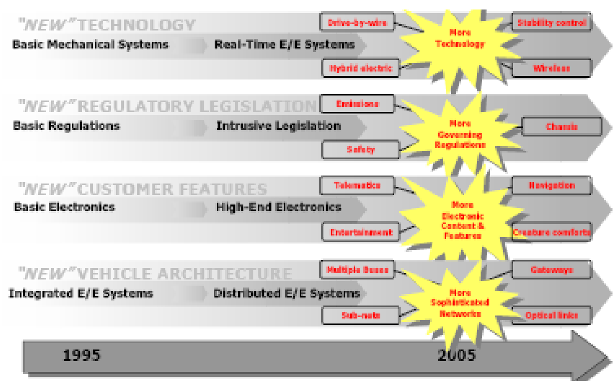


Figure 11. Increasing complexity of electronics in vehicles [29]

There are many opportunities for the KDD community to deal with the challenges in modern automotive sector. Some of these include:

- Standards development for achieving interoperability in sharing the engineering data in a geographically distributed organization (from vehicle design to after-sales);
- Development of common platforms (such as generic data models) for reusability to help in managing the data from a variety of vehicle lines;
- Use of multi-media and state-dependent test sequencing as a dynamic dimension to the existing static repair manuals;
- Develop technologies to mine not only the field data, but also metadata to create data models;
- Development of adaptive knowledge bases and advanced text analytics to deal with warranty, service, and customer relations data bases;
- Adaptation of text mining technologies to unstructured field data to build domain specific thesaurus/vocabulary/dictionaries;
- Promoting semantic web technologies (through domain ontologies development) for engineering data management;
- Widely deploying data-mining techniques (based on artificial intelligence, statistics, etc) for field data analysis tasks, such as sensor data pre-processing, clustering, fault classification, diagnostics and prognostics;
- Use of artificial intelligence techniques to develop advanced logistics and parts management systems;
- Technologies to aid in automotive internet activities (e.g., web analytics), where main challenges include data consistency & quality, as well as flexibility to meet the reporting needs of a company [33].

## 6 Appendix

Below is the list of acronyms that appeared in this paper.

*KDD: Knowledge Discovery and Data-mining*

*XML: eXtensible Markup Language*

*RDF: Resource Description Framework*

*CRISP-DM: Cross-Industry Standard Process for Data-Mining*

*PMML: Predictive Model Markup Language*

*SOAP: Simple Object Access Protocol*

*UDDI: Universal Description Discovery and Integartion*

*OLE BD-DM: Object Linking and Embedding for Data Bases for Data Mining*

*SGML: Standard Generalized Markup Languag*

*HTML: HyperText Markup Language*

*OWL: Web Ontology Language*

*W3C: World Wide Web Consortium*

*SPARQL: Simple Protocol And RDF Query Language*

*CAD: Computer Aided Design*

*FAR: Federal Aviation Regulation*

*NOTAM: NOTices to AirMen*

*AD: Airworthiness Directive*

*AECMA: Association Europeene des Constructeurs de Material Aerospatial (European Association of Aerospace Industries)*

*ATA: Air Transportation Association*

*DARPA: Defense Advanced Research Projects Agency*

*NASA: National Aeronautics and Space Administration*

*OASIS: Organization for the Advancement of Structured Information Standards*

*SAS: Statistical Analysis System*

*SPSS: Statistical Package for the Social Sciences*

*GATE: General Architecture for Text Engineering*

*SOM: Self-Organizing Maps*

# 7 References


[1] L. Huang, and Y. L. Murphey, "Text Mining with Applications to Engineering Diagnostics", IEA/AIE 2006, LNAI 4031, 1309-1317.

[2] A. Hotho, A. Nurnberger, and G. Paab, "A Brief Survey of Text Mining", LVD Forum – GLDV Journal for Computational Linguistics and Language Technology, 2005, 20(1): 19-62.

[3] K. Yamanishi, J. Takeuchi, and Y. Maruyama. "Data Mining for Security", NEC Technical Journal, 2003, Vol. 56, No.12.

[4] I. H, Witten, "Adaptive Text Mining: Inferring Structure from Sequences" J. of Discrete Algorithms, Vol. 0, 1-23, 0000.

[5] J. Wallace and T. Cermack., "Text Mining Warranty and Call Center Data: Early Warning of Product Quality Awareness", SUGI 29, 2004. Proceedings.

[6] B. D. Ville, "Text Mining with "Holographic" Decision Tree Ensembles" Proceedings of the Thirty-first Annual SAS Users Group International Conference, 2006

[7] B. D. Ville, "Ubiquitous Scoring of 1000+ Warranty Categories Using Predictive Rules Derived from Text", SAS Global Forum 2007.

[8] S. Ahmed, S. Kim, and K. M. Wallace, "A Methodology for Creating Ontologies for Engineering Design" ASME Design, Theory and Methodology Conference, 2005.

[9] J. Jarmulak, S. Craw, and R. Rowe, "Using Case-Base Data to Learn Adaptation Knowledge for Design", Proceedings of the Seventeenth IJCAI Conference, 2001, pp.1011-1016.

[10] I. Stock, and M. Waber, "Authoring Technical Documentation using a Generic Document Model", Proceedings of the SIGDOC'06.

[11] F. Servant, "Semantic Web Technologies in Technical Automotive Documentation", OWLED 2007.

[12] D. Macrini, D, A. Maggiore, and A. Santagata, "Semantic Web in the Automotive Industry: A Case Study", CEUR-WS 2005, Vol. 166.

[13] A. Saxena, "Knowledge-based Architecture for Integrated Condition Based Maintenance of Engineering Systems", PhD Thesis, Georgia Institute of Technology, 2007.

[14] J. Khan, and M. Hardas, "Hierarchical Course Knowledge Representation Using Ontologies", Proceedings of the IICAI, 2007

[15] L. Dey, M. Abulaish, and Jahiruddin, "Relation Extraction from Texts for Efficient Access to Information – Application to Bio-Medical Text Processing", Proceedings of the IICAI, 2007.

[16] W. Friedrich, "ARVIKA Augmented Reality for Development, Production, and Service", Proceedings of the International Symposium on ISMAR, 2002.

[17] M. Goebel, and L. Gruenwald, "A Survey of Data Mining and Knowledge Discovery Software Tools", SIGKDD Explorations, 1999, Vol. 1, Issue.1.

[18] B. Fortuna, M. Grobelnik, and D. Mladenic, "OntoGen: Semi-automatic Ontology Editor", HCI International 2007.

[19] Venkataraman et al. "Knowledge Stores for Interactive Diagnostics", US Patent: US20060106796.

[20] J. Keller, V. Beuer, and W. Kwedlo, "Application of Data-Mining and Knowledge Discovery in Automotive Data Engineering", PKDD 2000, LNAI 1910. pp.464-469.

[21] Cios, K. J., and Kurgan, L. Trends in Data Mining and Knowledge Discovery. Book: Knowledge Discovery in Advanced Information Systems, Springer 2002.

[22] Gersten, W., Wirth, R., and Arndt, D. Predictive Modeling in Automotive Direct Marketing: Tools, Experiences and Open Issues. Proceedings of the KDD 2000.

[23] U. M. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "Knowledge Discovery and Data Mining: Towards a Unifying Framework", In Knowledge Discovery and Data Mining, 1996, pp. 82-88.

[24] G. Creese, "Duo Mining: Combining Data and Text Mining. DM Review Online", September, 2004.

[25] OASIS Automotive Repair Information TC http://www.oasis-open.org/committes/tc_home.php?wg_abbrev=autorepair

[26] OWL Web Ontology Language Use Cases and Requirements: http://w3.org/TR/webont-req/#usecase-designdoc

[27] WIDE: Semantic Web-based Information Management and Knowledge Sharing for Innovative Product Design and Engineering. http://www.ist-wide.info/

[28] F. Servant, Renault, "Semantic Web Use Cases and Case Studies", W3C, May 2007.

[29] J. E. Hilger, E. Ford, and M. Flaherty, "Diagnostic Challenges in the Automotive Workshop", SAE Convergence, 2004.

[30] http://www.boeing.com/phantom/match_ct/iis.html

[31] S. Parthasarathy, "Data Mining at the Crossroads: Successes, Failures and Learning from Them", Proceedings of the KDD, 2007.

[32] T. A. Montgomery, Ford Motor Company, "Text Mining on a Budget: Reduce, Reuse, Recycle", Michigan Leadership Summit on Business Intelligence and Advanced Analytics: The New ROI – What is your Return on Information?, March 2007.

[33] M. Notte, "Deploying Web Analytics in a Pan-European Context", Toyota Motor Europe.

[34] T. Nsukawa, and T. Nagano, "Text Analysis and Knowledge Mining System", IBM Systems Journal. Vol. 40, No.4, 2001.