Spend analysis systems

S. Maržić, P. Krneta and M. Pavlić
Ris d.o.o., Kastav, Croatia
{sanjin.marzic, perino.krneta, mile.pavlic}@ris.hr

Abstract – This paper introduces spend analysis, spend management and supplier management field. It describes the need for a spend analysis system, its' benefits and savings which such system can help to achieve. We present flowchart of spend analysis process in form of a dataflow diagram. We present main issues in development of a spend analysis software, and give an overview of current spend analysis systems.

I. INTRODUCTION

Usage of computers in business, development of information systems, development of programming languages and databases made businesses more efficient. Reporting is faster, richer in content, more accurate, many tasks are now automated. All of these technologies participated and are still participating in collecting large amounts of various (structured) data. There is data of company's core business, procurement, retail, human resources, revenues, spend. This data contains useful information and knowledge but to extract it one has to make various *ad-hoc* analyses over large volumes of data. Such analyses aren't feasible on transactional systems because these systems are optimized for daily work which consists of simple, predefined operations and predefined reporting. This problem was solved with data warehouses – data was periodically exported from transactional system into another database, optimized for complex analyses. In data warehouse data was additionally enriched (fulfilled) by data from other (external) sources to achieve better analyses.

During the 1980s in the USA a new need emerged – spend analysis [1]. This analysis had different requirements which couldn't be fully satisfied by data warehousing. In this analysis commodity and supplier data had to be treated differently. Spend analysis' goal is to reduce the company's expense by current expense analysis, supplier analysis, commodity analysis. Expenses are also reduced by centralizing procurement and supplier rationalization.

Spend management is a way for a company to control and optimize its expenses; it includes cutting operational costs as well as other expenses [2]. Spend management gives a company control over the money so that all business needs (manufacturing, delivering of goods and services) are satisfied in the most cost efficient way. Profit of a company is the difference between total incomes and total expenses, so profit can be augmented either by raising the income side or by reduction of expenses. In (current) period of financial crisis reduction of expenses is a common approach because the reduction of 1$ on expenses equals to raising incomes for 3$ [2].

II. SPEND ANALYSIS

Spend analysis is organization of expenses data (mostly procurement) and usage of supplier hierarchy and hierarchy of goods and services combined with amounts spent to:

(1) Find categories of expenses

(2) Find opportunities for strategic sourcing by gathering procurement requirements and supplier rationalization

(3) Reduce expenses with higher discounts (with contracts, volume discounts).

Finding categories of expenses implies categorizing goods and services (expenses) so that one can answer questions such as what was total amount spent on computer monitors. Transactional systems usually contain only data on specific purchase (e.g. specific model of computer monitor), but to answer mentioned question one should collect data of all purchases. For this kind of analysis a hierarchy of goods and services is required. This hierarchy will, for specific model of computer monitor, define that it is a monitor of specific size (e.g. 17' monitor), an LCD monitor, hardware element etc. Using such hierarchy analysis is possible inside and between categories, e.g. which category contains most expenses, how large is price range of products or services of same category. Such analysis can be used to detect corruption which can be manifested as big difference in price for products in same category. One of the existing hierarchies for categorization of products and services is UNSPSC (United Nations Standard Products and Services Code) [3]. This hierarchy was developed by United Nations and Duns & Bradstreet. In this hierarchy every product and service has a unique 8-digit code in format SS-FF-CC-DD where SS is 2-digit representation of segment, FF is family, CC class and DD commodity. Important features of UNSPSC are (1) simple drill-down and drill-up by using parts of code, (2) consistency which ensures that every element is defined in exactly one point in hierarchy, (3) completeness because all products and services that can be traded are included in hierarchy and (4) market tracking meaning that hierarchy is maintained as new products and services appear in the market.

Gathering procurement requirements enables enlargement of single procurement orders and reduction of total number of orders (by connecting smaller orders to the same supplier into larger ones or connecting orders for

similar products or services). With larger orders company can get better procurement conditions and lower prices. **Supplier rationalization** implies creating supplier hierarchy. With this hierarchy expenses and procurement data can be categorized and company can define which suppliers are preferred and which are not. Supplier hierarchy defines which suppliers (that appear in transactional system) are parts of the same company (e.g. offices of the same company or multiple records for the same company) or are connected in some other way. With supplier hierarchy the real expenses towards some supplier can be determined (in contrast to expense towards an office of a supplier). By combining supplier hierarchy with products and services hierarchy questions like "how many suppliers are supplying some product category" and "how is the expense in some product category distributed over suppliers" could be answered. With this data, available **expenses can be reduced with higher discounts** by reducing the number of suppliers for a product category. Thereby the expense towards remaining suppliers would enlarge which creates better negotiating position for getting higher discounts. By detecting undesirable (not preferred) suppliers costs with them can be lowered and transferred to other suppliers (which can lower risks and raise quality).

A. Who needs spend analysis?

Not all companies have the need for spend analysis and there are several reasons for that: spend analysis software is usually very expensive and therefore not applicable for small companies and such companies don't even have the need for spend analysis since their procurement is often centralized and not very large in volume, so it can be analyzed without software. Companies that can benefit from spend analysis are the ones that are physically dislocated, without central procurement, with multiple procurement centers. The fact that some company is physically dislocated is not enough for it to need spend analysis: e.g. a retailing company with multiple supermarkets won't be interested in spend management software because, although it has multiple supermarkets, it has central procurement and all procurement data at one place. On the other hand city administration could be interested in spend analysis because a city usually has several communal companies which are financed from the same city budget, but have independent procurement. Since each company is autonomous in procurement substantial savings can be done by gathering the procurement requests. Also, corruption can be detected by comparing the prices for similar products.

B. Savings

A question that can be asked is: what kind of savings can be expected after implementing spend analysis and in which areas of business?

Table 1 displays possible savings (in percentage) by categories. Such savings are possible with commitment to advanced purchasing methods (e.g. e-sourcing, e-auctions) [4]. All savings are achieved in one of three ways [1]:

- Strategic sourcing – finding the strategic suppliers and reduction of supplier number to achieve discounts and raise quality (quality is raised by choosing the right supplier)
- Reduction of maverick spend by choosing (contracting) suppliers and improvement of business processes (e.g. internal guidelines for procurement)
- Avoiding of spend leakage by controlling suppliers' compliance to contract terms.

TABLE I. POSSIBLE SAVINGS WITH COMMITMENT TO ADVANCED PURCHASING [4]

Category	Saving
Raw materials	2-5%
Packaging	10-20%
Indirect materials and services	10-20%
Information technology	15-30%
Professional services (including consulting, legal human resources)	8-15%
Capital projects	7-15%
Other indirects (nonproduction costs)	5-15%
Media/marketing/promotional items	10-20%
Logistics (rail, truck, barge, ocean, air)	7-15%

Third method is especially important because a part of savings that were negotiated and agreed to can be lost without control of compliance. All three points should be continuously repeated to achieve continuous savings.

Data in table 1 is relative to full amount spent per category, specific amounts can be estimated with amounts of total expenses for each category. Since savings are expressed as percentage, saving of 2% for raw materials can seem like an insignificant saving, but actually it is the largest saving for production business systems since raw materials in those systems are carrying the largest part of total expense. Larger percentage for other categories can now seem like a much smaller saving but, as figure 1 depicts, it is not always so.

Total raw materials cost	1,000,000,000
Saving 5%	50,000,000
Indirect expenses, services	500,000,000
Saving 15%	75,000,000

Figure 1. Example of possible savings [4]

C. Specific advangates of spend analysis

Spend analysis software enables its users an overview of expense data for the entire company and reduces the need for additional data processing regarding reports and trend analysis. Specific advantages:

- Visibility of all expenses of the company
- Substantial improvement of data accuracy and consistency
- Improvement of analyses quality over time
- Reduction of time required for creating ad-hoc reports from up to two weeks to almost momentarily
- Reduction of maverick spend for additional savings
- Ability to perform data cleaning and create ad-hoc reports without need for permanently employed data extraction experts

- Reduction of administrational expenses for processing and controlling of procurement orders
- Fast identification of savings opportunities
- Additional savings by selecting suppliers and larger orders
- Savings by identification of contracting opportunities
- Elimination of overpaying the supplier and guarantee of rebate realization
- Continuous improvement of compliance to contract terms and savings by tracking contracted and real prices
- Reduction of procurements' depending on IT experts for supplying the expense data
- Removal of disagreements about data quality by introducing the "single truth" for spend data.

III. SPEND ANALYSIS PROCES

Spend analysis, as a process, is similar to the data warehousing process – both processes are based on ETLA process (Extraction, Transformation, Loading and Analysis).

Data extraction is periodical fetching of data from the transactional system, as well as from other relevant data sources, to make analyses up to date. Data can be fetched from sources other than databases, like user files, internet, etc.

Data transformation is the most important part of the spend analysis process (besides the analysis itself). Quality of analysis step depends on quality of data transformation. In this step transactional data is being normalized (conversion of currencies, units of measure) and brought to the common format (e.g. putting all address data in one field with specific ordering of data elements or dividing data to multiple fields). Another important segment of data transformation step, which distinguishes it from standard transformation in data warehousing, is building the products and services hierarchy and supplier hierarchy. Supplier hierarchy means connecting naturally connected suppliers (e.g. branch offices with head office or duplicate entries) so that it is easy to determine total spend with some supplier and not with e.g. its branch offices. Products and services hierarchy is development of multilevel hierarchy which will allow analyzing spend on multiple levels. E.g. management will be interested in general spend distribution and IT department will only be interested in subcategory of IT equipment expenses and will analyze more closely its sub-categories. Duplicate entries are detected and merged, business data is being cleaned. Part of this step is also data enrichment: missing data is being filled and external sources are being used (e.g. stock market data, supplier financial data). Since there is no perfect classification algorithm (for classification of suppliers and products) there must be an option for user to correct the results of data transformation. Data of users' interference to transformation results should be used to make future transformations more precise.

Data loading is loading the transformed data into a database optimized for spend analysis.

Data analysis is step where reports are created using the transformed data. Analysis, as with data warehouse, must enable the user to perform slicing & dicing, allow multidimensional and hierarchical view on data. Visualization is very important element of analysis; it should be able to work with large set of data efficiently.

Figure 2 depicts a model of the process described. Diagram was drawn according to DFC (data flow chart) method [5, 6]. Figure 3 depicts data transformation sub process.

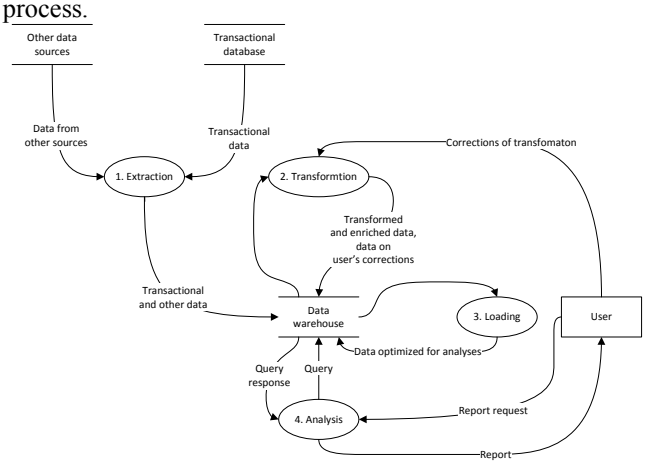

Figure 1. Spend analysis proces

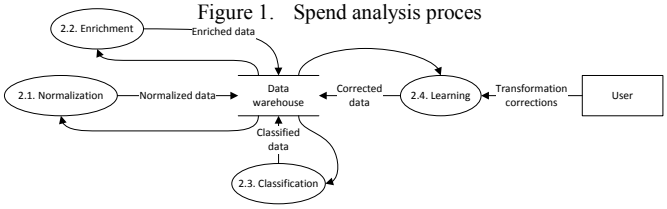

Figure 2. Data transformation process diagram

IV. SUPPLIER MANAGEMENT

Supplier management or supplier performance management is a process that consists of measuring, evaluating and controlling over suppliers' work, his business processes and habits with objective of reducing costs, lowering risk level and continuous progress [7].

Supplier management can be viewed as a business relation management for improvement of duplex understanding flow between a company and its suppliers, which especially concerns communication and negotiation about requirements and supply chain performance expectations [7]. Supplier management can include quality of delivered products and services, well-timed delivery, reliability of delivery (delivery of correct product in right amounts), adaptation to request changes, availability of required amounts of products or services when needed etc.

Supplier driven costs, or costs of poor quality (COPQ), can be divided into four groups [7]:

- Preventive expense – cost of actions that have to be performed to prevent appearance of bad products or services (work procedures, rules, planning, training)
- Evaluation expense – testing cost, evaluation or quality estimation cost
- Internal error expense – costs of low-grade products repairment (i.e. re doing the job)
- External expense – user support costs, dealing with customer complaints, loss of contracts (and reputation)

Low quality products and services costs are estimated to hold 10-25% of total sales, in that amount there is 25-75% cost produced by low supplier quality (2.5-18.75% of total sales) [7].

Low supplier quality costs are also affecting users by increasing price of company's products and services. Cost of low supplier quality rises with time needed to detect it. If problem is detected before the product is being used it can be rejected or returned to supplier which is the lowest cost; if, on the other hand, problem gets detected by the customer there are additional costs of customer service, complaint processing, returning the product, customer compensation and dealing with the returned product.

A good example of costs generated by using low quality product is cost of unusable X-ray image (table 2) [7]. Costs generated by one unusable X-ray image are paid time of technician and nurse who have to repeat the procedure, material expenses (new image, development materials, machine usage) and there is also a delay (and possible rescheduling) of the medical care that can't be given before usable X-ray image is taken. If such mistake happens five times per day and the cost of one error is evaluated to 33$ we get annual expense of 66,225$.

Information about total expense of low supplier quality is important because it helps prioritize such problems.

TABLE II. LOW SUPPLIER QUALITY EXAMPLE [7]

Element	Price
Technicians' time (30 minutes)	4$
Nurses' time (30 minutes)	9$
Film	7$
Development material	2$
Machine mainteanance	1$
Medical care delay	10$
Total price of one unusable image	**33$**
Annual expense (5 unusable imager per day, 365 days)	**66,225$**

V. SPEND MANAGEMENT SOFTWARE

First steps in spend analysis happened in 1980s, first spend management software appeared in early 2000s.

A. Properties of ideal spend analysis software [1]

In an ideal case spend analysis software would have the following elements:

1. Data selection and loading
2. Data cleaning, structuring and enrichment
3. Spend analysis
4. Knowledgebase management

Data selection and loading includes making a decision which data is going to be analyzed – what data is needed for spend analysis. With definition of necessary data, all data sources are defined [1]. Since data is fetched in a common database schema it should be brought to a compatible form, as data is arriving from heterogeneous sources. Data can be imported from different platforms, different databases, different file formats (Excel files, delimited files, etc.) [1].

Data cleaning, structuring and enrichment should have features:

1. Web oriented application with administrative and limited access – in cases where enrichment team is distributed each user should have access to specific products and services to prevent accidental (or unauthorized) change of data outside his/hers area of expertise
2. Data cleaning and enrichment – defined procedures for supplier data enrichment, classification of expenses (products, services), mechanism for conflict resolution. Expenses are classified by using rules (e.g. for supplier IBM expense is categorized as IT), these rules can be defined by an expert or by using computer learning algorithm. There can be a conflict set within set of rules (several rules with different outcome which can be activated in the same moment) so conflict resolution mechanism must be defined.
3. Defining and maintenance of hierarchies and dimensions – hierarchical schemes of products and services, supplier hierarchy, cost centers database are under continuous development so there must be one or more administrators for all these schemes, users should have limited authority over schemes and their parts to assure proper development of schemes.
4. Job scheduling – in large databases jobs can take a long time to complete, jobs should be well scheduled to assure the best system performance.
5. Import of end-user feedback – to improve system over time all feedback information should be collected to help in development of new rules based on collected information.

Spend analysis should have properties and functions:

1. Web-oriented application with administrative and limited privileges
2. Specialized data visualization tools
3. Reporting and dashboards

4. Sending feedback for changes in dimensions and hierarchies
5. Sending feedback for changes in classification
6. Immediate update of new data
7. *What-if* analyses

Knowledge base management means collecting data of user activity and to have an ability to learn upon collected data. System learns based on every action end user performs, so it should have ability to record every click of every user. Based on this information system can substantially improve its classification accuracy because it collects the data specific to user's industry. Knowledge database should have a lexicon to store data on three types of entities – concepts, instances and relations. Concepts are abstract terms (nouns, attributes, units of measure). Instances are descriptions of actual objects (e.g. "Dell 17' monitor"). Relations are structural and semantic links between concepts. Relations contain information on how concepts can be connected, which concepts are e.g. synonyms. Knowledge database is used in expense classification – the larger and more complete the lexicon is classification gets more accurate. Knowledge database management is an important task in spend analysis, there has to be a possibility to track changes in the database and to define a process for making changes, users who will have privilege of changing knowledge database.

In following sections some of the existing spend analysis applications are described.

B. Ariba

Ariba [8] is German software manufacturers founded in 1996, specialized for development of spend analysis software, contract management software, supplier management software. Their spend management application is "Ariba Spend Analysis". This application is used in software as a service model (SaaS). In 2012 Ariba was acquired by SAP.

Data imported into application is enriched with data from Dun & Bradstreet database (D&B). It is an American company that collects data on other companies to improve decision making, supplier chain management. Data is also enriched with data regarding ecology of business, company relations, and risks.

Properties of application that manufacturer points out is that it enables comparison with similar companies in the same industry; it enables predictive analyses and *what-if* analyses, classification (of products and services, suppliers) in predefined hierarchies or user defined hierarchies. Project manager is assigned by Ariba to every client.

C. Bravo Solution

Bravo Solution [9] is Italian spend management software manufacturer. The application "Bravo Solution Spend Management" is their application for spend management; as with Ariba, it is offered in SaaS model.

Properties of application that manufacturer points out is possibility to use multiple hierarchies to classify expense data, user can define their own hierarchies. Expenses are classified based on rules and classifiers trained by machine learning algorithm based on previous decisions. In supplier classification D&B database is being used as well as comparison to other available supplier attributes.

D. CVM Solutions

CVM Solutions [10] is supplier analysis software manufacturer. They are collecting supplier data so that suppliers themselves are supplying specific documents and reports, CVM Solutions are checking this data with public and closed sources and are continuously monitoring changes in supplier data. In this way they've collected a large volume of supplier data, so their service is supplier data cleaning (removing of multiple entries for same supplier, linking suppliers that are naturally connected), supplier data enrichment, analysis and control of company's suppliers and warnings in risky situations.

E. Applicability to Croatian (regional) circumstances

In Croatia and surrounding states there is not a great number of large companies which have the need for spend analysis, so spend analysis and management is not a developed field in the region. Above mentioned software is partially applicable to local companies because of several reasons: (1) D&B database and CVM Solutions database contains none or very small volume of data on regional companies so this data can't be used in supplier data cleaning; cleaning can only be done by record comparison, (2) mentioned solutions don't mention if their application interface is translated to Croatian language (or any of regional languages). In product classification some applications state that stop-words are used (words that don't contribute to the meaning of the statement and are ignored in text analysis, e.g. conjunctions, prepositions, articles), English stop-words aren't applicable to other languages so for Croatian and regional languages a research should be conducted to retrieve a list of such words. Another aspect that manufacturers aren't stating publicly is their software price. Since their customers are large corporations (they reference AT&T, Dell, Cisco, McDonalds, Mitsubishi), software price can be high because for these companies it will pay off with savings they will achieve. In region there is not very large number of large multinational companies, so their savings won't be so large; this is why software price will be a significant factor.

Existing software solutions are applicable to regional companies, but it can't be expected to give as good results as they achieve in "English speaking" companies. In Croatia a project MaxyTransparency [11] is started to develop a software solution that will be adjusted to language(s) and specific need of regional market.

VI. SPEND ANALYSIS ISSUES

In previous sections we've introduced spend analysis, requirements and causes of appearance, properties of an ideal software application and examples of existing

solutions. In development of a generic software application for spend analysis (or introduction of spend analysis to a business system) there are several issues (question) that have to be solved in different phases of the process. These issues and questions are displayed in table 3.

TABLE III. SPEND ANALYSIS ISSUES PER PHASES

Data extraction	
Refresh period	How often data extraction from transactional system has to be performed (weekly, monthly, once every quartal, yearly)? How recent analysis have to be? Should some data be imported more often?
Size of extraction	What data will be extracted/imported? Only new data or entire database will be refreshed? How to distinguish old and new data? How to recognize what data is already imported, but was changed meanwhile? Are some changes on transactional system required to optimize extractions? What data is allowed to be extracted? (important issue for demonstration of spend analysis power/ROI because real possibilities cannot be shown with missing data)
External sources	Which external sources will be used? How often? What is their price? How reliable/relevant are that sources?
Data transformation	
Product and services classification	How will products and services be classified? Which (how many) schemes will be used? How is the classification going to be performed? Which algorithms will be used?
Supplier hierarchy	How is the supplier hierarchy going to be built? How to detect/remove multiple records for the same supplier?
Data enrichment	What data will be used for data enrichment? (reliability of sources, relevance of sources, cost of this data)
Learning	How to improve future classification based on users interaction? Which algorithms will be used? How to train initial model (on what data)?
Errors	If an error is detected in data should source be corrected as well?
Data loading	
Schema	What schema (model) will be used? OLAP cubes are precalculating "answers to all questions", but for price of long refresh time and cubes are unable to give the *what-if* scenario analysis. Relational models, such as dimensional model, are faster in refresh, but queries take longer to execute.

Additional issues are present for spend analysis applications in regional circumstances. Some of reasons are language specificities (e.g. multiple word forms), writing specificities (special characters, often misuse of letters *č* and *ć*, replacements of *č* and *ć* with *c* or *đ* with *dj*). Spend analysis for companies that primary use English language can use results of numerous researches in natural language processing while in Croatian language and regional languages this research it is not possible because lack of interest for development of this area. There are no researches or tagged corpora for supplier (or products) matching model training.

Special problem is also to persuade companies that, if they are cohesive units, it is good for them to connect in unified procurement system to gather suppliers and products into a single supplier management system and usage e-supply (e.g. mutual e-auctions).

VII. CONCLUSION

In this paper we've introduced spend analysis. We've displayed expected results of spend analysis application usage and also given examples of losses, which happen without noticing, that occur when spend analysis is not applied. Process of spend analysis and supplier management is described. We've given list of features that software application for spend management should have. Also, we've analyzed usability of existing spend analysis applications on regional circumstances and identified issues that should be resolved in development of a spend analysis system.

Implementation of spend analysis system is recommended in every larger organization, public institution or private corporation, especially with large competition, lowering of income and market crisis. Spend management system can produce better results in combination with e-procurement system and supplier management system. E-procurement system helps the procurement sector; in combination with results from spend management system enables the procurement to focus on products and suppliers where most of savings can be achieved. Supplier management system holds an up-to-date supplier database, their products and services which are being used by both spend management system and e-procurement system. Such database can be created by joining several consumers and suppliers in single network: consumers can use it as an e-commerce system and suppliers as a platform for placement of their products and services.

One of suitable areas for implementation of these three systems is public sector (e.g. Republic of Croatia) and we hope that such system will be developed in near future.

ACKNOWLEDGMENT

The research has been conducted under the project "Extending the information system development methodology with artificial intelligence methods" (reference number 13.13.1.2.01.) supported by University of Rijeka (Croatia).

LITERATURE

[1] Pandit, K., Marmanis, H. (2008). *Spend analysis – the window into strategic sourcing*. J. Ross Publishing

[2] Wikipedia (2013). *Spend Management*, http://en.wikipedia.org/wiki/Spend_management

[3] Granada Research (2001). *Using the UNSPSC – United Nations standard products and services code*. Granada Research

[4] Rudzki, R. A., Smock, D. A., Katzorke, M., Stewart, S. J. (2006). *Straight to the bottom line*. J. Ross Publishing

[5] Pavlić, M. (2011). *Informacijski sustavi*. Školska knjiga, Zagreb

[6] Pavlić, M. (2011). *Oblikovanje baza podataka*, Odjel za informatiku Sveučilišta u Rijeci

[7] Gordon, S. R. (2008). *Supplier evaluation and performance excellence*. J. Ross Publishing

[8] Ariba http://www.ariba.com/

[9] Bravo Solution https://www.bravosolution.com/

[10] CVM Solutions http://www.cvmsolutions.com/

[11] Ris (2012) *Maxy Transparency: project documentation*. Ris d.o.o.