SPEND ANALYSIS GLOSSARY
# TABLE OF CONTENTS

<p>| ACTUAL SPEND | 4 | DATA SCRUBBING | 12 |
| ADDRESSABLE SPEND | 4 | DATA WAREHOUSE | 12 |
| AI | 4 | DATABASE | 12 |
| AP | 4 | DIRECT SPEND | 12 |
| BI | 5 | DIRTY DATA | 12 |
| BIG DATA | 5 | DIVERSITY SPEND | 13 |
| BLACK BOX | 5 | DOUBLE-BYTE (ASIAN) CHARACTERS | 13 |
| CATEGORY MANAGEMENT | 6 | eCLASS | 13 |
| CLASSIFICATION ALGORITHMS | 6 | ePROCUREMENT | 13 |
| CONTROL TOTALS | 6 | ERP | 13 |
| COST MANAGEMENT | 6 | FIREWALL | 14 |
| COST REDUCTION | 6 | FTP | 14 |
| COST SAVINGS | 6 | GL | 14 |
| CPO | 7 | HISTORIC SPEND | 14 |
| DASHBOARD | 7 | INDIRECT SPEND | 14 |
| DATA CLASSIFICATION | 7 | INITIAL CLASSIFICATION | 15 |
| DATA CLEANSING | 8 | INVOICE DATA FILE | 15 |
| DATA COLLECTION | 8 | KPI | 16 |
| DATA ENRICHMENT | 9 | LEVERAGING SPEND | 16 |
| DATA EXTRACT | 10 | MACHINE LEARNING | 16 |
| DATA FEED | 10 | MARKET INDEX | 16 |
| DATA MAPPING | 10 | MARKET INTELLIGENCE | 16 |
| DATA QUALITY | 10 | MARKET TRENDS | 16 |
| DATA REFRESH | 11 | MASTER DATA MANAGEMENT | 17 |
| DATA SET | 11 | MATERIAL GROUPS | 17 |</p>
<table>
<thead>
<tr>
<th>Glossary Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAVERICK SPEND</td>
<td>17</td>
</tr>
<tr>
<td>MRO</td>
<td>17</td>
</tr>
<tr>
<td>MRP</td>
<td>17</td>
</tr>
<tr>
<td>NAICS</td>
<td>17</td>
</tr>
<tr>
<td>NONTRADITIONAL SPEND</td>
<td>18</td>
</tr>
<tr>
<td>OFFSHORE</td>
<td>18</td>
</tr>
<tr>
<td>OLAP</td>
<td>18</td>
</tr>
<tr>
<td>OUTSOURCE</td>
<td>18</td>
</tr>
<tr>
<td>PARENT/CHILD RELATIONSHIP</td>
<td>18</td>
</tr>
<tr>
<td>PARETO PRINCIPLE</td>
<td>18</td>
</tr>
<tr>
<td>pCARD</td>
<td>19</td>
</tr>
<tr>
<td>PERFORMANCE MANAGEMENT</td>
<td>19</td>
</tr>
<tr>
<td>PO</td>
<td>19</td>
</tr>
<tr>
<td>POC</td>
<td>20</td>
</tr>
<tr>
<td>PREFERRED SUPPLIER</td>
<td>20</td>
</tr>
<tr>
<td>PROCUREMENT</td>
<td>20</td>
</tr>
<tr>
<td>PURCHASING</td>
<td>21</td>
</tr>
<tr>
<td>REAL-TIME DATA PROCESSING</td>
<td>21</td>
</tr>
<tr>
<td>RISK MANAGEMENT</td>
<td>21</td>
</tr>
<tr>
<td>RFI</td>
<td>21</td>
</tr>
<tr>
<td>RFP</td>
<td>21</td>
</tr>
<tr>
<td>RFQ</td>
<td>22</td>
</tr>
<tr>
<td>ROI</td>
<td>22</td>
</tr>
<tr>
<td>SAAS</td>
<td>22</td>
</tr>
<tr>
<td>SCM</td>
<td>22</td>
</tr>
<tr>
<td>SIC</td>
<td>22</td>
</tr>
<tr>
<td>SOURCEABLE SPEND</td>
<td>23</td>
</tr>
<tr>
<td>SOW</td>
<td>23</td>
</tr>
<tr>
<td>SPEND</td>
<td>23</td>
</tr>
<tr>
<td>SPEND AGGREGATION</td>
<td>23</td>
</tr>
<tr>
<td>SPEND ANALYSIS</td>
<td>24</td>
</tr>
<tr>
<td>SPEND ANALYTICS</td>
<td>24</td>
</tr>
<tr>
<td>SPEND BENCHMARKING</td>
<td>24</td>
</tr>
<tr>
<td>SPEND CATEGORIZATION</td>
<td>24</td>
</tr>
<tr>
<td>SPEND CLASSIFICATION</td>
<td>24</td>
</tr>
<tr>
<td>SPEND CUBE</td>
<td>25</td>
</tr>
<tr>
<td>SPEND DATA</td>
<td>25</td>
</tr>
<tr>
<td>SPEND DATA MANAGEMENT</td>
<td>25</td>
</tr>
<tr>
<td>SPEND INTELLIGENCE</td>
<td>25</td>
</tr>
<tr>
<td>SPEND MANAGEMENT</td>
<td>26</td>
</tr>
<tr>
<td>SPEND TAXONOMY</td>
<td>26</td>
</tr>
<tr>
<td>SPEND UNDER MANAGEMENT</td>
<td>26</td>
</tr>
<tr>
<td>SPEND VISIBILITY</td>
<td>26</td>
</tr>
<tr>
<td>SUPPLIER COMPLIANCE</td>
<td>27</td>
</tr>
<tr>
<td>SUPPLIER DATA FILE</td>
<td>27</td>
</tr>
<tr>
<td>SUPPLIER DIVERSITY</td>
<td>28</td>
</tr>
<tr>
<td>SUPPLIER ENRICHMENT</td>
<td>28</td>
</tr>
<tr>
<td>SUPPLIER GROUPING</td>
<td>28</td>
</tr>
<tr>
<td>SUPPLIER NORMALIZATION</td>
<td>28</td>
</tr>
<tr>
<td>SUPPLY CHAIN</td>
<td>29</td>
</tr>
<tr>
<td>TAIL SPEND MANAGEMENT</td>
<td>29</td>
</tr>
<tr>
<td>UNSPSC</td>
<td>29</td>
</tr>
</tbody>
</table>
ACTUAL SPEND

Most companies define **actual spend** as the sum of all payments made through accounts payable (AP) system(s). There may be other systems handling spends that are not recorded in AP, but most companies view AP as the corporate “checkbook.” What is a company’s actual spend? Let’s look at this question from a couple of different perspectives. If we were to ask our clients this question, many would likely respond with: “Our total spend is X, but our actual spend is Y.” So why the difference? The term “actual spend” is often used to describe the **amount of spend that is sourceable**. Let’s say a company has an annual spend of $5 billion. When you remove the intercompany spends and the categories that are not sourceable, you may end up with addressable spends around $2.5 to $3 billion. This figure represents what most sourcing professionals believe is the actual spend, or the amount that the sourcing team can proactively address.

Strategic-sourcing consulting organizations provide another perspective on this question. These consulting organizations are hired to identify “waves” of savings opportunities by category. From their perspective, the actual spend is the **total of spends for the categories in one of the waves**.

ADDRESSABLE SPEND

Addressable spend is the portion of sourceable spend that is currently addressable. For example, shipping is typically sourceable, but the client may be under a contract with FedEx that cannot be negotiated until the end of the next year. Therefore, it is not currently addressable.

AI

**Artificial intelligence (AI)** within the domain of spend analysis is a software application whose purpose is to aid in the execution of spend analysis tasks such as classification, cleansing, etc. The “intelligence” emerges from the fact that AI is designed to correctly solve problems that it has never seen before. For example, spend data often comprises unstructured text containing misspelled words, cryptic descriptions, acronyms, dates and measurements. This is the type of problem data that AI is designed to correctly analyze and classify. AI is a very desirable component of a larger spend analysis solution because it provides the ability to analyze large volumes of data quickly.

AP

**Accounts payable (AP)** is money owed by a business to its suppliers and shown on its balance sheet as a liability. Also, the business department or division that is responsible for making payments owed by the company to its suppliers and other creditors.

Accounts Payable is one of the common data elements used in typical spend analysis, among others:

1. General ledger
2. Purchase order
3. Supplier data files
4. Invoice data files
5. pCard

**BI**

Business intelligence (BI) is a standard term often heard and perhaps most liberally interpreted. Business intelligence is the identification, extraction and analysis of business data to provide an overview of business operations. Spend analysis falls under the umbrella of business intelligence—it involves identifying, extracting and analyzing a company’s spend data to provide an overview and analysis of the company’s spend. The analytics derived from a spend analysis would be a subset of the overall business intelligence.

**BIG DATA**

While the term “big data” has become a buzzword, spend professionals don’t agree on its meaning. Many websites disagree on what quantity makes “big data” or avoid quantifying the term. But in reality, the amount of digital data generally available is exploding, and the term applies itself well to today’s spend analysis initiatives. Internal spend data, along with external, third-party spend and supplier-enrichment data is increasing exponentially. Project implementations can total billions of dollars in spend, with tens of millions of line-item transactions. For spend analysis, this data isn’t just captured in a master data management module, but it is cleansed, classified, analyzed, then refreshed and analyzed all over again.

**BLACK BOX**

In spend analysis terms, a black box is a database with predetermined rules for cleansing and classifying spend. This may sound like a great idea, and it was, initially. Prior to the black box, spend analysis was done manually. It was labor-intensive and prone to human error. As databases and computers evolved, companies and programmers realized they could “automate” spend analysis by programming rules in a database that would automatically cleanse and classify spend. For example, rules can be programmed to always change “FedEx,” “Fedx” and “Fed Ex” to “Federal Express.”

**CATEGORY MANAGEMENT**

In terms of spend analysis and sourcing, category management is used by the procurement and purchasing function for sourcing, acquisition, and supplier management. A “category” is the logical grouping of spend on similar items, such as office supplies or IT hardware. Grouping the similar items together in categories, or “sourcing groups,” makes it easier for the procurement department to ensure that goods and services are being sourced and purchased from the same preferred supplier, that contract terms are being adhered to, and that total spend by category and supplier can be calculated.

Small companies may only attribute their spend to a small number of categories. Larger companies, however, use hundreds of categories, thereby getting a deeper analysis of spending patterns and trends as well as sourcing and savings opportunities.
CLASSIFICATION ALGORITHMS

A classification algorithm, in the context of spend analysis, is a series of steps that one follows in order to take unclassified spend data and assign a classification code to each line item according to one's taxonomy, e.g., identifying a category as UNSPSC 10101501. There are many different algorithms, and each algorithm has its strengths and weaknesses. A good algorithm will be flexible enough to accurately classify items with few mistakes and in a short period of time.

CONTROL TOTALS

A control total is a total for one or more numeric fields that appears in every record. One example is the total payment for a group of suppliers.

COST MANAGEMENT

Cost management is the process of collecting, analyzing, evaluating, and reporting cost information for use in budgeting, estimating, forecasting, and monitoring costs. Cost management can refer to the costs associated with individual projects or an overall business. We refer to cost management of an overall business as it relates to spend management. Cost management is an integral part of a company's success. When properly implemented, cost management will provide reduced production costs for products and services, resulting in increased value for the customer.

COST REDUCTION

Cost reduction is the process used by companies to reduce their costs associated with developing, producing, manufacturing, and supplying their product lines. There are several ways to reduce costs:

- Using lean manufacturing principles
- Standardizing parts
- Reducing overhead costs
- Ensuring the quality of the product, thereby reducing what is known as the “quality cost”
- Managing and consolidating the supply chain
- Adhering to contracts, using preferred suppliers, and reducing maverick spend
- Performing regular spend analyses

COST SAVINGS

Cost savings is the amount of money a company saved on a product or service due to its cost reduction and cost management strategies.
**CPO**

A chief procurement officer (CPO) is an executive role within a company, generally reporting to the CEO or the president. Primary responsibilities are:

- Sourcing
- Procurement
- Procurement policies and procedures
- Supply management
- Supplier relations
- Contract management
- Compliance management
- Cost management
- Pricing negotiations
- Asset management
- Spend analysis

**DASHBOARD**

A dashboard is a customized user interface that organizes and presents information from multiple components into a unified display that is easy to read and interpret. It serves as a starting point from which a corporate executive can see the big picture before digging deeper into data. Spend analysis dashboards typically display:

- Spend by top suppliers
- Spend by category
- Spend by location
- Spend by business unit

**DATA CLASSIFICATION**

In order to easily and systematically access important or often-used files, structures are put in place to classify or sort information. When information is placed in expected locations, users can better track and use the data. In the case of spend data, data classification is essential to ensure proper identification of line-item transactions at the category level as well as the supplier level. Without accurate, in-depth line-item transaction classification, potential sourcing and savings opportunities may be missed and spend may not be fully understood.

Spend data classification requires all items to be assigned to a classification schema. A typical classification schema is UNSPSC; however many companies use customized classification structures.
There are three methods of data classification:

1. Manual data classification
2. Automated data classification
3. Combination of manual and automated data classification

**Manual data classification** involves the human interaction of a person (or team of people) manually reviewing line-item transactions and manually classifying them. Manual data classification is time-consuming and prone to human error and subjectivity, which leads to low accuracy levels. An often used example is Polaris. One person may classify Polaris as a pool cleaner; another person may classify Polaris as a heart monitor. Upon reviewing and analyzing your data, it may not be obvious that this item has been classified in two different categories. This will lead to inaccurate reporting on the true spend for a category, item and supplier, and may negatively impact contract negotiations. It can also affect budgets, forecasts, and inventory levels.

**Automated data classification** involves a computer program automatically classifying line-item transactions. The flexibility and robustness of the program will determine the level of accuracy for data classification. Can the program support multiple taxonomies? How robust are the rules that classify items? Can the program differentiate between a pool cleaner and a heart monitor by looking at description, supplier, price, sourcing category, quantity and buyer? Depending on the level of detail the program can support when classifying items, results can be accurately classified or they can be misclassified.

**Combination of manual and automated data classification** can be the best approach when used properly. A flexible master data management (MDM) tool that supports multiple, simultaneous classification taxonomies and allows robust rules to be applied at the item level, combined with a process of manually reviewing and classifying items that do not adhere to existing classification rules, will provide the most accurate results. This is true for the initial classification as well as for future data refreshes.

**DATA CLEANSING**

This is the process of reviewing data and cleaning it up with the goal of fixing incorrect, out-of-date, redundant, duplicate, incomplete or incorrectly formatted data. Similar to supplier normalization, data should have a consistent look and feel, making it easier to read and ensuring the accuracy of the analysis and reporting. In some industries, data cleansing is also known as data scrubbing.

**DATA COLLECTION**

Data collection describes a process of preparing and collecting data to record, to make decisions about important issues, or to pass on to others. For spend analysis, data collection describes the process of collecting and aggregating all spend-related data. Data collection should be completed on an ongoing basis to allow for data refreshes, thus ensuring the most relevant, up-to-date spend information. Data can be retrieved via an electronic data feed or through Excel spreadsheets.
Spend data can be found in:

- General ledger files
- Accounts payable files
- Supplier files
- Invoice files
- Purchase order files
- Expense files
- e-Procurement files

Typical fields in a spend data file are:

- Supplier name
- Supplier code
- Supplier location
- Line-item description
- Material/category description
- GL account description
- Spend amount
- Source system

DATA ENRICHMENT

For spend analysis purposes, data enrichment refers to any external third-party data that will supplement or enrich existing spend data with the goal of making better financial decisions. Data enrichment also enhances the value, quality, and integrity of data, making it more complete, consistent, and up-to-date. Both free and paid enrichment data are available from trusted sources; however, free enrichment data may not provide the right level of detail. Data enrichment for spend analysis can include the following:

- Supplier parent/child information
- Supplier diversity information
- Producer price indexes
- Commodity pricing
- Global market data
- Freight information
- Supplier demographic information
- Supplier financial and operational information
DATA EXTRACT

A data extract comprises data retrieved from various sources for further processing or storage. For the purposes of spend analysis, acceptable forms of data for a data extract are:

- Excel data file with consistent internal formatting
- Text file
- File with data linkages maintained

Forms of data that are not acceptable for a spend analysis data extract are:

- File with data linkages not maintained
- Comma-separated file with nothing to separate data fields that may contain a comma
- Excel data file with multiple formats embedded
- Excel report file

DATA FEED

A data feed is an interface through which a client application, such as a web browser, can stream data—typically in real time. The data can be of any type, including video, audio, text, etc. Within the context of spend analysis, data feeds can be the typical spend data files such as AP and PO as well as commodity indices, financial data, etc.

DATA MAPPING

Data mapping is the process of “mapping” distinct data models together and linking them. When collecting large amounts of spend data from many different systems, data mapping is used to determine the relationship between the different spend data sources and the destination where that data is used. This allows the disparate data to be merged into one final database for cleansing and classifying.

For example, spend transactions that have been classified to UNSPSC can be mapped to custom categories. The UNSPSC codes for colorants, dyes, pigments, color compounds and dispersions, and polymer masterbatches may all be mapped to the “Resin-Colorants” category.

DATA QUALITY

Data quality is the fitness of the data to serve its purpose in a given context. For spend analysis, the following aspects should be considered when determining the quality of spend data:

- Accuracy of the data
- Completeness of the data
- Timeliness of the data
- Age of the data
Data quality is crucial to the trustworthiness of spend data for reporting, analysis and the ability to make business decisions based on reporting and analysis.

**DATA REFRESH**

Refreshing the data displayed on a web page is as easy as clicking the refresh or reload button. In terms of spend data, however, there are a number of methods to refresh your data, and they can provide dramatically different results. First and foremost, the results of a spend data refresh will depend on the amount of time and effort spent upfront on the initial spend data classification. Another factor in a spend data refresh is whether new data elements are being introduced into your data, in which case detailed data cleansing and data classification will need to take place. Similar to the initial spend data classification, there are three methods of data refresh available:

1. Manual data refreshes
2. Automated data refreshes
3. Combination of manual and automated data refreshes

**Manual data refreshes** involve the human interaction of a person (or team of people) manually reviewing line-item transactions and manually classifying them. Basically, a manual data refresh involves performing an initial cleansing and classification all over again. Manual data refreshes are time-consuming and prone to human error and subjectivity, which leads to low accuracy levels. An often-used example is Polaris. One person may classify Polaris as a pool cleaner; another person may classify Polaris as a heart monitor.

**Automated data refreshes** involve programs such as natural language processing or machine learning to automatically classify line-item transactions. The most common issue with automated data classification is that the same rule will always apply to the same item. Therefore, if Polaris has been identified as a pool cleaner in the initial cleansing and classification, it will always be classified as a pool cleaner. This also leads to low accuracy levels.

Used properly, the combination of automated and manual data refreshes can be the best approach when new data elements have been introduced. A flexible master data management tool that supports multiple, simultaneous classification taxonomies and customized item classification is required, as well as an in-depth knowledge of the business to ensure items are being classified accurately.

**DATA SET**

A data set is a collection of related sets of information. For example, a data set for spend analysis might contain all travel and expense spend data.
DATA SCRUBBING

Also known as data cleansing, data scrubbing is the process of reviewing all data and cleansing it with the goal of fixing incorrect, out-of-date, redundant, duplicate, incomplete or incorrectly formatted data. Similar to supplier normalization, you want your data to have a consistent look and feel, making it easier to read and ensuring the accuracy of the analysis and reporting.

DATA WAREHOUSE

A data warehouse is a database with a rigid architecture, generally used for offline applications. Data warehouses have typically been poor mediums for “closing the loop”—getting data back to the original source environment. The databases simply maintain a copy of the information from the original source.

DATABASE

A database (DB) is an application that manages structured sets of data organized for convenient access and to be manipulated and updated. Databases are organized by fields, records and files. A field is a single piece of information, such as a field for data purchased. A record is one complete set of fields, such as all the fields related to one line-item transaction. A file is a collection of records, such as all purchase order records. A database can be thought of as an electronic filing system.

DIRECT SPEND

Direct spend is spend that goes directly into the development and manufacture of a product. It refers to the costs of goods and services that are directly incorporated into a product being manufactured, or Cost of Goods Sold (COGS).

DIRTY DATA

Dirty data is any data that contains errors. Dirty data can be caused by a number of factors including duplicate records, incomplete or outdated data, improper storage methods and incorrect data mapping. The following data is considered dirty:

- Incorrect data: Does not conform to the field’s values; for example, a buyer name in a price field
- Inaccurate data: May be in the correct fields and correctly formatted, although the data itself is incorrect
- Business rule violations: Violates business rules such as having an expiration date prior to a purchase date
- Inconsistent data: Has not been cleansed or normalized
- Incorrectly punctuated or misspelled data
- Incomplete data: Contains missing or null values; for example, a line-item transaction missing the description or the supplier name
• Duplicate data
• Data without generalized formatting

DIVERSITY SPEND

To demonstrate their core values, many companies have implemented programs to support the growth of diverse suppliers. Since the early 1950s, companies that supply the federal government with goods and services have been required by law to have a supplier diversity program or initiative. In order to measure and report on the success of these programs, companies must be able to track the amount of spend with diverse suppliers.

DOUBLE-BYTE (ASIAN) CHARACTERS

The Chinese, Japanese and Korean (also referred to as CJK) writing systems use 16-bit (2-byte) character sets. English, by contrast, is a single-byte language. As an alphabetic language, each letter in the English alphabet occupies a single byte in computer memory. The CJK languages are syllabic languages, and each syllable occupies two bytes in computer memory. The maximum number of characters that can be represented with 1 byte (English) is 256 characters, whereas 2 bytes (CJK) can represent up to 65,536 characters.

eCLASS

eClass is the de facto classification standard within the German energy industry but competes with UNSPSC in the United States.

ePROCUREMENT

The term eProcurement arose when electronic procurement systems entered the market, allowing buyers and procurement departments to electronically manage the entire purchasing process at their companies. Controlling who could buy what, from which supplier and at which price—as well as approving purchases, categorizing purchases and saving purchase history via an electronic program—eliminated the use of paper purchase orders and manual approval processes. eProcurement refers to all purchases made by a company through an eProcurement system both from catalogs uploaded to the system and through suppliers’ eProcurement websites. eProcurement can also refer to electronic purchases made through online marketplaces. Purchases made from a supplier’s website are not considered eProcurement since they are made outside the procurement system.

ERP

ERP stands for enterprise resource planning, which involves using an integrated system with access to a solid database. The basic platforms and software can be very expensive to implement and maintain, but the premise of an ERP is a good one. It gives insight into data at various levels, including accounting, human resources, supply-chain management, project management, customer-relationship management, and manufacturing. The end result is a platform that allows for an efficient flow of information about the various business transactions inside and outside the organization. Most spend analysis initiatives involve, but are not limited to, spend data from ERP systems.
FIREWALL

A firewall is a system designed to prevent unauthorized access to or from a private network and to protect its resources. Software that is installed “behind the firewall” indicates that it has been installed on a company’s private network and is protected and maintained by its IT department. This software is only accessible on-site or via Virtual Private Network (VPN). Software-as-a-Service (SaaS) is not installed behind a company’s firewall and is accessible anywhere via a browser. Security is maintained by the provider.

FTP

File Transfer Protocol (FTP) is a standard network protocol used to transfer files from one host to another host over a TCP-based network, such as the Internet. FTP is often used to transfer large spend data files.

GL

The general ledger (GL) is the main accounting record of a business that uses double-entry bookkeeping. The general ledger should include the date, description, and balance or total amount for each account. It is usually divided into at least seven main categories. These categories generally include assets, liabilities, owner’s equity, revenue, expenses, gains and losses.

The GL is one of the common data elements used in typical spend analysis, among others:

1. Purchase order (PO)
2. Supplier data files
3. Invoice data files
4. pCard
5. Accounts payable (AP)

HISTORIC SPEND

Historic spend is often used to track, review and analyze past spending practices. By tracking trends, companies aim to improve current and future spending practices.

INDIRECT SPEND

Spend that is not attributable to product development and manufacturing is called indirect spend. Examples include the lubricants used to maintain the machines that produce a product or the lights in a factory. Indirect spend typically shows up as “operating costs” or SG&A (selling, general and administrative) costs.
INITIAL CLASSIFICATION

At the beginning of a spend analysis project, an organization will perform an initial classification by collecting typical spend data (such as AP, PO, GL, pCard, supplier data and invoice data) then classifying that data into sourcing groups. Category or sourcing managers will then break down the data with the goals of driving sourcing opportunities and gaining a better understanding of the organization’s spend and suppliers.

This initial “snapshot” approach typically evolves into the need for ongoing data refreshes to provide better reporting and insights. As capabilities to manage ongoing data are put into place, this quickly generates demand for quarterly to monthly refresh cycles.

INVOICE DATA FILE

Examples of fields in an invoice data file include:

- Source system ID
- Invoice line number
- Description
- Invoice number
- PO line number
- PO description
- Amount USD
- Quantity US
- Estimated savings USD
- Accounting date
- UNSPSC code
- Commodity ID
- Supplier ID
- Supplier location ID
- Unit of measure
- Cost center ID
- Requester user ID
- Account ID
- Invoice date
- Paid date
- Invoice type
- Contract ID
- Diversity status
- Payment terms
- PO amount
- Price variance
- Standard cost variance
- Original currency code
- Original amount

The invoice data file is one of the common data elements used in typical spend analysis, among others:

1. General ledger (GL)
2. Purchase order (PO)
3. Supplier data files
4. pCard
5. Accounts payable (AP)
KPI

Key performance indicators (KPIs) are specific, measurable, and actionable values that are identified as important in determining the success or failure of an activity. KPIs vary by activity, but they can also vary for the same activity, depending on what is important at the time. With spend analysis, KPIs may focus on contract adherence one year and savings the next.

LEVERAGING SPEND

Leveraging spend refers to using the total spend you have spent with a supplier to negotiate better contract deals. See also spend aggregation.

MACHINE LEARNING

The term machine learning is a broad one that refers to a number of algorithms with the objective of evolving behaviors based on observed characteristics within a set of examples. Within the context of spend analysis, machine learning has applications within classification, cleansing, etc., as it can recognize patterns within the underlying data, and these apply to future decisions.

MARKET INDEX

A market index is an aggregate value produced by combining several stocks or other investment vehicles together and expressing their total values against a base value from a specific date. Market indexes are intended to represent an entire stock market and thus track the market’s changes over time.

Index values are useful for investors to track changes in market values over long periods of time. For example, the widely used Standard and Poor’s 500 Index is computed by combining 500 large-cap U.S. stocks together into one index value. Investors can track changes in the index’s value over time and use it as a benchmark against which to compare their own portfolio returns.

MARKET INTELLIGENCE

This area of research involves the collection and analysis of data on individual markets, providing a broad view of market sizing, share and segmentation. Data is typically gathered from sources such as company accounts, trade organizations, interviews with business contacts and consumer research. Market intelligence is not considered a component of business intelligence but rather a separate and independent research field.

MARKET TRENDS

Market trends are the large-scale movements of financial markets over time. Stock traders and analysts attempt to identify current market trends and predict how long they will last before moving in the opposite direction. These trends can strongly influence an organization’s decisions to buy or sell investments. The phrase “bull market” refers to an upward market trend, while the phrase “bear market” refers to a downward trend.
MASTER DATA MANAGEMENT

Master data management (MDM) refers to the processes and technologies used in creating a secure and centralized data environment to represent a “single version of the truth” for an organization. The master data management system doesn’t just hold data; it collects data from numerous disparate sources, and it functions as the accepted system of record for use across all software applications, business units, and user groups.

Examples of master data include:

- Customer data (name, contact details, DOB, customer classification)
- Locality data (physical address, postal address, geographical data)
- Product data (item number, bill of materials, product codes)
- Employee data (employee number, role, placement in organizational structure)
- Partner data (partner name, classification)

MATERIAL GROUPS

A group of materials with the same attributes or purpose, used mainly with SAP software.

MAVERICK SPEND

Also known as “rogue spend” or “noncompliant spend,” this refers to the spend associated with any item or service that is purchased outside the preferred supplier list and/or outside the preferred process or system. According to a recent research report by American Express and AT Kearney, maverick spend accounts for up to $433 billion at European companies.

MRO

Maintenance, repair, and operations (MRO) is often considered the same as indirect spend, or at least a part of it.

MRP

Materials requirement planning (MRP) is a computerized ordering and scheduling system that calculates and maintains an optimum manufacturing plan based on master production schedules, sales forecasts, inventory status, open orders and bills of material (BOM). If properly implemented, it will improve cash flow and increase profitability. MRP will provide the ability to be proactive rather than reactive in the management of inventory levels and material flow.

NAICS

The North American Industry Classification System (NAICS, pronounced “nakes”) has replaced the outdated Standard Industrial Classification (SIC) system. It is an industry classification system, not a product classification system, and therefore not well suited for procurement or sourcing.
NONTRADITIONAL SPEND

Nontraditional spend refers to the spend for categories that are not traditionally sourced, managed, or analyzed by the procurement department. These categories may include:

- Health benefits
- Advertising
- Travel
- Fleet services

OFFSHORE

Offshore simply refers to work done for a company by people in another country. In terms of spend analysis, offshore specifically refers to the cleansing and classification of spend data completed by offshore resources.

OLAP

OLAP is an acronym for online analytical processing. OLAP performs multidimensional analysis of business data and provides the capability for complex calculations, trend analysis, and sophisticated data modeling. OLAP provides data in a numeric view rather than a graphical view.

OUTSOURCE

Not to be confused with offshore, outsource refers to any task, operation, job, or process that could be performed by employees within an organization but is instead contracted to a third party. These functions can be performed by the third party on-site or off-site.

PARENT/CHILD RELATIONSHIP

Parent/child relationship describes how children (subsidiaries and divisions) relate to a parent company. A linkage is the relationship between different companies or specific sites within a corporate family. This occurs when one business location has financial and legal responsibility for another business location.

There are two types of linkage relationships:

- Branch or division to headquarters
- Subsidiary to parent

PARETO PRINCIPLE

Also known as the 80-20 rule, the Pareto Principle states that for many events, roughly 80 percent of the effects come from 20 percent of the causes. In terms of spend analysis, the Pareto Principle would state that 80 percent of an organization’s spend comes from 20 percent of its suppliers.
pCARD

A purchasing card or procurement card (pCard) is a company credit card issued to employees to procure goods and services. Controls such as single-purchase dollar limits, monthly spend limits, supplier restrictions or category code limits can be implemented.

pCard data is one of the common data elements used in typical spend analysis, among others:

1. General ledger (GL)
2. Purchase order (PO)
3. Supplier data files
4. Invoice data files
5. Accounts payable (AP)

PERFORMANCE MANAGEMENT

Setting limits to spend, setting limits to the number of suppliers per category, and adhering to contract terms are all part of performance management. Processes need to be put in place to ensure that the savings identified through spend analysis are being monitored and managed dynamically. The goal is to manage and maintain the performance of all savings and sourcing initiatives. Also known as “spend performance management” and “supplier performance management.”

PO

A purchase order (PO) is a written sales contract issued by a buyer to a seller detailing the exact merchandise or services to be purchased. It will specify payment terms, delivery dates, item identification, quantities, shipping terms and all other obligations and conditions.

PO data is one of the common data elements used in typical spend analysis, among others:

1. General ledger (GL)
2. Supplier data files
3. Invoice data files
4. pCard
5. Accounts payable (AP)
POC

A proof-of-concept (POC) is generally the evidence that a product or technology is viable and capable of solving an organization’s particular problem. A proof-of-concept is usually small and may or may not be complete.

In terms of spend analysis, organizations often require that a spend analysis provider complete a POC before making a final decision. A POC usually includes one month, one quarter, or one year of spend, depending on the size of the company. It may also include just one region or one business unit. Companies are usually looking for the accuracy of the results from the POC to determine if the provider is a viable option. The spend analysis provider and the organization should agree how they will determine the level of accuracy before the POC.

PREFERRED SUPPLIER

A supplier is designated as “preferred” by the procurement department due to many factors such as pricing and contract terms. Purchases from suppliers that are not preferred are considered maverick spend.

PROCUREMENT

Traditionally, procurement simply involved the acquisition of a good or service. However, the role and value of the procurement department have become more strategic and more aligned with a company’s bottom line.

The “seven rights” rules (7 Rs) that procurement should follow when buying products and services are:

1. Right price
2. Right quantity
3. Right quality
4. Right time
5. Right place
6. Right source
7. Right service

To manage the procurement process and the supply base efficiently and effectively, procurement should focus on:

Identifying opportunities:

- Evaluate and select a diverse and globally competitive supply base
- Monitor and interpret supply markets and trends and their impact on company strategies
- Identify new potential suppliers and introduce contingency plans
- Act as the primary contact with suppliers
Managing internal operations:

- Manage of procurement staff
- Develop and maintain policies and processes
- Identify, introduce and leverage appropriate technology and systems
- Define the procurement strategy and structure
- Develop plans and measures

**PURCHASING**

Purchasing is the specific function associated with the actual buying of goods and services from suppliers.

**REAL-TIME DATA PROCESSING**

Within the context of spend analysis, real-time data processing is an automated step of cleansing and classifying data in order to supply the results quickly to a user. The “real-time” component enables consumers to view accurate data quickly without the delay included in typical cleansing and classification jobs. Services put the data into the processing pipeline as soon as it is available and generate reports for the user.

**RISK MANAGEMENT**

For spend analysis purposes, risk management refers to the identification, assessment and prioritization of the supply chain risks impacting an organization’s spend.

**RFI**

A request for information (RFI) is a buyer’s call for written information about the capabilities of suppliers. RFIs are commonly used on major procurement projects where a requirement could potentially be met through several alternate means. An RFI, however, is not an invitation to bid, is not binding on either the buyer or sellers, and may or may not lead to an RFP or RFQ.

**RFP**

A request for proposal (RFP) is a call for suppliers or vendors to submit proposals for an upcoming project or engagement. For example, a new business might request proposals for the hardware, software and user training required to establish and integrate a new system. An RFP is issued in the early stages of the procurement process and is meant to clearly identify the risks and possible benefits associated with a project.
RFQ

A request for quotation (RFQ) is the process by which a company invites suppliers or vendors to bid on a specific project or engagement. RFQs typically include detailed specifications on products or services in order to ensure that suppliers are bidding on the same products/services. Suppliers are required to enter their competing bids by a specific date, at which point the company may choose to hold further rounds of negotiations before ultimately awarding its business to the supplier with the best offer.

Similar requests include a request for information (RFI) and a request for proposal (RFP).

ROI

Return on investment (ROI) can be viewed in two different ways for the purposes of spend analysis. The first view is the basic ratio of money gained or lost, positive or negative, as compared to the money spent. Spend analysis can be used to determine the ROI for specific suppliers, such as diversity suppliers.

The second view refers to ROI for a spend analysis initiative. How much money was saved after a spend analysis compared to how much money was spent to conduct the spend analysis? The calculation of ROI for a spend analysis initiative could be the amount of savings discovered minus the cost of the solution divided by the cost of the solution.

SAAS

Software-as-a-Service (SaaS) is a software delivery model in which software and associated data are centrally hosted by the provider and available to users via the Internet. Also referred to as “on demand,” “web based,” “Internet based,” or even “cloud computing,” SaaS has become the industry standard for software delivery.

SaaS has many advantages—it is available anywhere, at any time and from any machine with Internet access. No upgrades are necessary, as the provider pushes out new releases and bug fixes seamlessly to all customers at the same time. The onus of installing, maintaining and securing third-party software is taken off the hands of the internal IT department and placed back on the provider.

SCM

Supply chain management (SCM) is the design, planning, execution, control and monitoring of an organization’s supply chain activities. The objectives are to create net value, build a competitive infrastructure, leverage worldwide logistics, synchronize supply with demand and measure performance globally.

SIC

The Standard Industrial Classification (SIC) is a U.S. government system for classifying industries by a four-digit code. Established in 1937, it is being replaced by the six-digit North American Industry Classification System (NAICS), which was released in 1997.
SOURCEABLE SPEND

Sourceable spend is the portion of a company’s total spend for which savings opportunities can be explored. Nonsourceable categories would include taxes and government payments, employee benefit programs (401k, payroll), intercompany spend, etc.

SOURCING

Sourcing refers to a number of proactive procurement practices aimed at strategically identifying, evaluating and engaging suppliers to find the most cost-efficient goods and services.

SOW

The terms scope of work (SOW) and statement of work (SOW) have the same meaning and are often used interchangeably. An SOW is a legal, formal document that identifies and describes the work to be performed or the services to be provided. SOWs should be completed before work begins. Any changes should be detailed in a change order.

In addition to pricing, the SOW should answer the following:

- **Who?** Define roles and identify who will perform specific tasks.
- **What?** Define goods; services and exceptions; and exact scope and size of SOW, including quantities, tasks, methodologies, work activities, deliverables and requirements.
- **Where?** Determine locations to perform services, deliver goods and exceptions.
- **When?** Start and stop dates, milestones, project plans and timeframes, frequency and timelines.
- **Why?** Include the background and purpose of the SOW.

SPEND

The most common definition for spend is “to pay out or expend money.” It’s traditionally considered a verb; however in the spend analysis industry, spend is used as both a verb and a noun. One might ask, “How much do you spend annually?” or, “What is your annual spend?”

SPEND AGGREGATION

This refers to combining buying power in specific categories in order to secure the best pricing and service agreements from suppliers. This can also refer to leveraging spend. If a company can identify it’s spending $10M with a particular supplier, it is more likely to negotiate better contract terms than if it’s only spending $1M with that supplier.
SPEND ANALYSIS

Spend analysis is the act of analyzing a company’s spend in an effort to reduce costs and increase savings. Spend analysis can be broken up into three separate areas:

1. **Data Collection**: This involves the entire process of collecting, cleansing, classifying and analyzing spend with the purpose of reducing procurement costs, improving efficiency, and monitoring compliance.

2. **Data Maintenance**: Companies categorize and evaluate their spend at the category or item level in order to identify areas that may offer savings opportunities.

3. **Reporting and Analysis**: Spend analysis should be an ongoing effort rather than a one-time initiative and should be an in-depth analysis of spend, including the who, what, when, where, and why. Whom do you spend with, what do you spend on, when do you spend it, where do you spend it and why do you spend it?

SPEND ANALYTICS

Simply stated, the analytics derived specifically from a spend analysis and used for reporting and further analysis.

SPEND BENCHMARKING

Spend can be benchmarked **internally**, such as between departments, categories and business units. It can be benchmarked to the **market** (determining how much other comparable companies spend on this same category), or it can be benchmarked against **historic spend** (determining how much a company spent in the last year).

SPEND CATEGORIZATION

Synonymous with spend classification, this refers to the process of grouping spend into categories so that an organization can easily manage its spend. The number of categories can run into the hundreds, depending on the company and the industry. Categories can be structured using taxonomies such as UNSPSC, eClass, or company proprietary taxonomies. Spend can be categorized at a high level, such as “office supplies,” or it can be detailed, such as “legal-size paper.”

SPEND CLASSIFICATION

Synonymous with spend categorization, this refers to the process of grouping spend into categories so that an organization can easily manage its spend. The number of categories can run into the hundreds, depending on the company and the industry. Categories can be structured using taxonomies such as UNSPSC, eClass, or company proprietary taxonomies. Spend can be categorized at a high level, such as “office supplies,” or it can be detailed, such as “legal-size paper.”
SPEND CUBE

A spend cube aggregates data in each level of each dimension in a given OLAP schema. The business intelligence industry uses the word “cube” because it best describes the resulting data. Also known as an OLAP cube or a data cube.

SPEND DATA

This refers to any and all data that assists with management of a company’s spend. This typically includes:

- Item number
- Price
- Unit of measure
- Category
- Supplier

It can also include much more robust information such as:

- Producer price indexes
- Commodity pricing
- Freight information
- Parent/child linkage
- Global insight market data

SPEND DATA MANAGEMENT

This refers to the actual management, monitoring and maintenance of any and all data that relates to a company’s spend. This can include the entire process of collecting, cleansing, and classifying the data.

SPEND INTELLIGENCE

Spend intelligence refers to the analysis performed on spend data after it has been cleansed and classified, enabling better sourcing decisions and providing insight into savings opportunities.
SPEND MANAGEMENT

Spend management takes a spend analysis and turns it into actionable savings and sourcing. Spend management incorporates all aspects of spend analysis, spend categorization/spend classification, spend data management and spend performance management. Spend management is an ongoing process and a company-wide initiative to control and optimize a company’s spend, effectively improving its bottom line. The spend management process should include the following steps:

1. Define objectives, goals and benchmarks
2. Obtain data from all sources and consolidate
3. Ensure data accuracy and completeness
4. Analyze the data
5. Turn the data into business intelligence, make smart decisions and take strategic action that saves money and increases profits
6. Repeat

SPEND TAXONOMY

Spend taxonomy is the common classification taxonomy used throughout an organization to provide consistent reporting. Most organizations support more than one classification taxonomy, such as UNSPSC, eClass, or a homegrown classification structure, and struggle to map those to one standardized internal taxonomy.

SPEND UNDER MANAGEMENT

The percentage of a company’s entire spend that procurement manages is considered the percentage of “spend under management.” The lower the percentage of “spend under management,” the lower the potential savings and sourcing opportunities. All spend from all departments should be managed under procurement to ensure that a company is sourcing and spending efficiently.

SPEND VISIBILITY

Visibility in the spend management area refers to the ability of an organization to gain a comprehensive view of the metrics that drive improved cost savings, process efficiency and supply-chain performance. Specifically, spend visibility enables analysis of past spend to make smart future business decisions.
SUPPLIER COMPLIANCE

Supplier compliance refers to a buyer organization’s set of rules by which all suppliers must conduct business. These rules help companies streamline and standardize their internal procedures for dealing with suppliers. Suppliers will often make adjustments to their operations to react to their customer’s requests. Compliance rules may include:

- Payment terms
- Contract terms
- Shipping terms
- Environmental reporting
- Responsible sourcing

SUPPLIER DATA FILE

Examples of fields in a supplier data file include:

- Source
- Vendor/supplier number
- Name
- Street name/address
- City
- State/province
- Country
- Postal code
- Telephone
- PO box
- PO box postal code

The supplier data file is one of the common data elements used in typical spend analysis, among others:

1. General ledger (GL)
2. Purchase order (PO)
3. Invoice data files
4. pCard
5. Accounts payable (AP)
SUPPLIER DIVERSITY

Many companies require purchasing from suppliers designated as diverse. Examples include:

- **Small Business (SBA):** Must meet the numerical size standards as defined in the Small Business Size Regulations, 13 CFR 121.
- **Minority Business Enterprise (MBE):** The business is at least 51 percent owned by such individuals or, in the case of a publicly owned business, at least 51 percent of the stock is owned by one or more such individuals. Further, these minority group members control the management and daily operations.
- **Women Business Enterprise (WBE):** The business is at least 51 percent owned by such individuals or, in the case of a publicly owned business, at least 51 percent of the stock is owned by one or more such individuals. Further, the management and daily operations are controlled by the woman or women members.
- **Veteran-Owned Business (VOB):** Must be at least 51 percent owned and controlled by a U.S. veteran or veterans possessing a discharge other than dishonorable.

SUPPLIER ENRICHMENT

This refers to the process of enriching, improving or adding complementary data to existing supplier data, which is beneficial to an organization’s spend management initiatives. Additional data such as supplier risk, supplier diversity, producer price indexes, commodity pricing and freight information helps organizations make better, more informed decisions.

SUPPLIER GROUPING

This refers to the grouping together of all entities, divisions or parent/child relationships of a particular supplier. For example, Kinko’s would be grouped with Federal Express. This ensures an organization is viewing the combined total spend with a supplier, allowing it to negotiate better contract terms.

SUPPLIER NORMALIZATION

As with any data, supplier names need to be entered consistently in order to produce accurate reports and analytics. When supplier names aren’t entered consistently, they need to be “normalized.” In other words, an organization must determine that the word “incorporated” will always be identified as “Inc.,” that “limited liability company” will always be identified as “LLC,” and that all variations of “FedEx” will be identified as “Federal Express.” Once these decisions have been made, rules can be written (similar to “find and replace”) to ensure consistency with supplier names.
SUPPLY CHAIN

A supply chain is a system of organizations, people, technology, activities, information and resources involved in moving a product or service from supplier to customer. Supply chain activities transform natural resources, raw materials, and components into a finished product that is delivered to the end customer.

TAIL SPEND MANAGEMENT

Tail spend comes from the Pareto Principle and refers to the 20 percent of spend that will come from 80 percent of suppliers. This 20 percent of spend is typically the most difficult spend to categorize and analyze. Tail spend contains a large number of disparate suppliers and categories, and it may contain a large amount of maverick spend. Since this last 20 percent doesn’t offer “quick wins” for savings, it is often ignored.

UNSPSC

UNSPSC is an acronym for the United Nations Standard Products and Services Code. It is a hierarchical classification system used to classify all products and services on five levels. The system is intended to efficiently identify a company’s spend by using a standardized classification code. A company can then easily pull up records on products and suppliers. UNSPSC is a popular classification and is used in ERP systems, data classification systems, business intelligence and in spend management. It is not, however, the best option for sourcing items since sourcing groups don’t always adhere to UNSPSC classification.