ABOUT OPERA SOLUTIONS:
Opera Solutions applies advanced analytics to drive front-line productivity gains for industry and government. With 220+ data scientists among its 650 employees, we are an internationally recognized leader in machine learning. We combine this scientific expertise with the deep industry knowledge of our professionals to create our Signal Hub™ solutions. These provide new enterprise-wide capabilities to continually extract valuable predictive and descriptive information from Big Data flows. We also create and host delivery of applications built on Signal Hub solutions. We serve leading organizations in Financial Services, Healthcare, Government, Supply Chain, Marketing, and other selected sectors.

Opera Solutions’ Supply Chain Solutions group focuses on driving new levels of productivity and savings gains at the front lines of procurement operations, marrying Opera’s Signal Hub solutions and proprietary spend optimization technology with the deep domain knowledge of our professionals. The applications we build on our Signal Hub solutions put advanced machine intelligence at the service of procurement professionals, bringing automated savings alerts and insights, critical market information, and analytics continually to their desktops. For more information, email SupplyChainSolutions@operasolutions.com.
Authors' Note

Although examples in this document are based on situations encountered during actual spend analysis efforts, all data has been disguised and the name fabricated.
Executive Summary

Almost any attempt by an organization to analyze spending patterns is likely to be fruitful, especially if there hasn’t been a serious prior attempt. It is easy to find thousands of breathless testimonials about a particular product or method – independent of the quality of the product or method – because almost any product or method will find savings if a spend visibility initiative has never been launched before. “In the land of the blind, the one-eyed man is king.”

This simple fact has confused end-user organizations and analysts for many years. In fact, it has convinced most spend visibility vendors (and most analysts) that spend visibility is a fundamentally simple process of mapping Accounts Payable spend, and then drilling for dollars. This is why many spend analysis products have remained largely unchanged for years; there is no perceived need to do anything "more".

What is not so obvious is that this initial burst of savings is short-lived; and that many of the "quick saves" that result are unsustainable. The key question is what to do next; in other words, how to implement a true strategic spend visibility initiative that will return value and keep returning value over time. There are too many spend visibility products that are lying unused or on the shelf, after the first burst of excitement has passed; and too many organizations who are tired of hearing a spend visibility message that has no further relevance to them.

Advanced organizations understand that strategic spend visibility is much more than building a simple AP cube, and that analysis of spend requires deep thinking on many dimensions, along with many different analysis cubes. The question is, how is this to be accomplished? Although much of the strategic spend analysis "lore" has been locked up inside consulting organizations, this is starting to change. We hope that this Guide will help to promulgate some of the key ideas around strategic spend visibility, and ideally point organizations toward strategies that can result in sustainable savings through a continuous succession of intelligent spend control initiatives.
Table of Contents

List of Figures......................................................................................................................... 7
Background and Need .................................................................................................................. 9
  Substantial Savings Opportunities ......................................................................................... 10
  The Spend Analysis Value Curve ......................................................................................... 10
  A Need for Strategy ................................................................................................................ 12
Basic Process .............................................................................................................................. 12
  01 Get Executive Support .................................................................................................. 12
  02 Understand Existing Capabilities and Gaps .................................................................. 13
  Understand your users and potential users ..................................................................... 20
Identify the Data Sources .......................................................................................................... 22
  03 Identify the Data Sources ............................................................................................. 22
  04 Identify the Raw Data .................................................................................................. 23
Defined the Schema .................................................................................................................. 26
  Why Do We Need A “Master” Cube? Remember the Basic Question .............................. 26
Merging Similar Datasets .......................................................................................................... 32
Merging Disparate Datasets ...................................................................................................... 33
Special Case Merging ............................................................................................................... 35
Address Critical Issues ............................................................................................................. 37
Don’t Get Confused By Language ......................................................................................... 40
Check Your Control Totals ...................................................................................................... 40
Family the Vendors .................................................................................................................. 43
Define the Relationships .......................................................................................................... 45
Define the Commodity Category Structure .......................................................................... 47
Define the Starting Rules for Classification and Mapping ...................................................... 51
Use the Secret Sauce ............................................................................................................... 51
Don’t Forget the 80/20 Rule .................................................................................................... 54
Remember, AI is Not Enough ................................................................................................. 58
Don’t Fall for the Accuracy Myth ............................................................................................. 60
Avoid the Common Tactical Traps ............................................................................................. 61
ERP is Not the Answer ............................................................................................................. 61
A Centralized Warehouse Doesn’t Solve the Analysis Problem ............................................ 62
A Common Taxonomy Only Creates New Problems .............................................................. 64
Crosstabs Aren’t the Answer Either ....................................................................................... 66
Define the Cubes ...................................................................................................................... 70
Define the Base Cubes .............................................................................................................. 70
Build Enriched Cubes .............................................................................................................. 71
  Channel Compliance ........................................................................................................... 71
  Spend Under Management ................................................................................................. 72
Diversity .................................................................................................................................. 74
Risk ....................................................................................................................................... 75
Compliance ............................................................................................................................... 76
Analyze, Assess, Report, Decide ............................................................................................. 77
Find Opportunities in the Data ................................................................................................. 77
Start with Immediate Payback Opportunities ....................................................................... 80
Best-Price Contracts ................................................................. 81
Substitutions .................................................................................. 82
Service Contracts ........................................................................... 84
Duplicate Payments ........................................................................ 88
Move on to Low-Hanging Fruit Opportunities ....................................... 88
Off-Contract Spend .......................................................................... 89
Maverick Spend ............................................................................... 91
Significantly Above Index Spend ......................................................... 94
Then Get Creative ............................................................................ 95
Start with a Total Cost of Ownership .............................................. 95
Focus on Often Ignored Cost Components ........................................ 96
Identify Underlying Cost Drivers ...................................................... 97
Control Demand ............................................................................ 98
Rationalize the SKUs ...................................................................... 100
Analyze Inventory .......................................................................... 102
Manage Energy .............................................................................. 102
Plan On Reporting As Part of the Process ....................................... 103
To Each Her Own ........................................................................... 104
C-Suite Executives .......................................................................... 105
CPO and Procurement Managers .................................................... 105
Analysts ........................................................................................ 109
10 Refresh, Repeat .......................................................................... 111
Best Practices ............................................................................... 114
Formulate a Plan ........................................................................... 114
Analyze Everything ......................................................................... 114
Build as Many Cubes (and Reports) as Possible ............................. 115
Drill, Drill, Drill ............................................................................ 117
Help Suppliers Understand What They Are Costing The Organization 119
Increase Spend Under Management .............................................. 119
Treat All Categories as Equal ........................................................ 120
Share .......................................................................................... 120
Create a Cube-Based Knowledge Store ........................................ 120
Measure and Report the Results ...................................................... 121
Train, Train, Train ......................................................................... 121
Actively Look for Strategic Opportunities ...................................... 121
Pair Spend Analysis with Decision Optimization .......................... 122
Don’t Be Afraid to Outsource Analysis ......................................... 123
Selecting the Right Platform ............................................................. 123
Great (E)TL Support ...................................................................... 123
Real-Time Changes ........................................................................ 124
Intuitive ........................................................................................ 124
Visual ........................................................................................... 125
Resource Light ............................................................................. 125
Competitive Cost ........................................................................... 125
Output to Formatted Reports ......................................................... 125
Performance ................................................................................. 125
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Successful Test Drive</td>
<td>126</td>
</tr>
<tr>
<td>Selecting the Right Architecture</td>
<td>128</td>
</tr>
<tr>
<td>Summary</td>
<td>129</td>
</tr>
<tr>
<td>References</td>
<td>132</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>134</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1 The "Classic" Spend Analysis Value Curve .......................................................... 10
Figure 2 The Modern Spend Analysis Value Curve .......................................................... 11
Figure 3 Changing Hierarchies ......................................................................................... 15
Figure 4 Changing Keys ..................................................................................................... 16
Figure 5 Creating Arbitrary Measures ............................................................................. 17
Figure 6 Building Advanced Reports ............................................................................... 18
Figure 7 Changing Data Formats ...................................................................................... 19
Figure 8 Expected Value for Ad-hoc Data Analysis ......................................................... 20
Figure 9 Organizational Data Sources for Spend Analysis ............................................. 22
Figure 10 AP Record Examples ........................................................................................ 24
Figure 11 Invoice Example ............................................................................................... 25
Figure 12 The "Who buys What from Whom" Cube ......................................................... 27
Figure 13 The Clustered Crosstab Report ........................................................................ 28
   Figure 13b A series of The Clustered Crosstab Report
   Figure 13c Detail from a block in a Clustered Crosstab Report
Figure 14 An AP Record Merge Example .......................................................................... 31
Figure 15 An AP and PO Record Merge Example .............................................................. 34
Figure 16 A Merged PO Report ......................................................................................... 36
Figure 16a Revised Merged PO Report 1 .......................................................................... 37
Figure 17 Different Names for the Same Purpose .............................................................. 40
Figure 18 Mapping Annual Report Spend to Cash in the System .................................... 41
Figure 19 A Time Series Analysis ..................................................................................... 42
Figure 20 A Year-Over-Year Spend Baseline .................................................................... 43
Figure 21 A Vendor Grouping Example ............................................................................ 44
Figure 22 Segmented Vendor Grouping Example ............................................................ 44
Figure 23 Duplicate Vendor Example ............................................................................... 45
Figure 24 Vendor Grouping Overview ............................................................................. 46
Figure 25 Different Organizational Views ....................................................................... 47
Figure 26 Mapping Rule Overlay, Visually Explained ...................................................... 56
Figure 27 A Mapping Rule Overlay Example .................................................................. 57
Figure 28 Mapping Percentages ....................................................................................... 58
Figure 29 Accounting Code Examples ............................................................................. 63
Figure 30 UNSPSC Example ............................................................................................ 65
Figure 31 Direct Mail Spend Example ................................................................................ 66
Figure 32 MMG Crosstab Commodity Spend Report ...................................................... 67
Figure 33 Vendor Spending By Organizational Unit Crosstab .......................................... 68
Figure 34 Quarterly Vendor Spending by Organizational Unit Crosstab .......................... 69
Figure 35 Facility by Contractor Rate Range Crosstab ...................................................... 69
Figure 36 Some Basic Spending Cubes ............................................................................. 70
Figure 37 A Sample Compliance Report .......................................................................... 72
Figure 38 A Combined Spend View ................................................................................ 73
Figure 39 Sample MWBE Report ..................................................................................... 74
Figure 40 Sample MWBE Goal by Operating Committee Member Report .................... 75
Background and Need

One need look no farther than Strategic Spend Visibility: [An] Untapped Potential for Cost Reduction[^6] for an explanation of why spend visibility is needed. Alternatively, consider the recent Global Supply Chain Study[^8] by IBM, based on face-to-face interviews with nearly 400 supply chain executives in 25 countries in which 70% of executives said that “their number one challenge is overwhelming and fragmented data, as well as a lack of ability to make sense out of the information”.

The sober reality is that procurement data is a mess at most organizations that have not implemented a proper spend visibility solution. Consider the case of the company that bought office supplies online using punch-out in their PO (purchase order) system, as described in Opportunity Analysis: The Challenge is Having Accurate and Usable Spend Information[^11], first published on Sourcing Innovation. The vendor processes the orders, ships the items, and presents the invoices for payment – invoices which are approved in the PO system. When the data was analyzed:

- item numbers in the PO system did not match items in the contract for 20% of items purchased,
- there was no way to tell what percentage of spending was for items on-contract vs. items off-contract,
- agreed-to price changes were not tracked and the impact could not be calculated, and
- demand had shifted to the point where 75% of contracted items were no longer being bought and 40% of the top purchases were without contract pricing.

This was just one category, yielding incremental savings of 5% when the problems were fixed. Now imagine what this means when translated across all categories.

What might not be so clear is why a spend visibility implementation guide such as this one is required. The reality is that if spend visibility is not approached in a strategic and methodical manner, then the resulting implementation will likely be only tactical in nature, with just a short term benefit. Some savings will be achieved, as per the discussion in Strategic Spend Visibility, but the reality is that those savings will start to trickle off quickly, often within twelve months (or less). However, if a well planned spend visibility program is followed, savings will materialize year after year ... and those savings will be substantial.
**Substantial Savings Opportunities**

Spend Visibility is one of the most powerful sourcing tools in a buyer’s arsenal. According to a recent Aberdeen Group research study, an average company can save 11% with a modern spend analysis system[1]. That’s 11 million of savings on every 100 million of addressable spend. In fact, a buyer will often save 5% simply by using the information to consolidate their supply base in a cost effective manner[2]. Furthermore, if a strategic approach is taken to spend visibility and analysis efforts, addressable spend and savings opportunities are multiplied and the organization can achieve hundreds of millions of dollars in savings, like PPG who saw savings of over 500 million within two years[3]. Plus, there’s always a chance the savings opportunity could be even bigger than the organization expects. As Chris Sawchuk, Procurement Practice Leader at Hackett Associates says, *it’s hard to say what your opportunities are without gaining visibility*[13].

**The Spend Analysis Value Curve**

An initial spend visibility project will always yield savings because almost any attempt to analyze spend will identify some spending that is not at contracted rates, is off contract, or is prime for consolidation and/or negotiated contracts. Each of these efforts will yield savings. Furthermore, a simple analysis of spend by supplier will yield opportunities for supply base consolidation, lack of adherence to “best-price” clauses, and volume-based discounts and rebates. Opportunity after opportunity will present itself during the first six-to-eighteen months, until the top-N reports are exhausted. Then, if a strategic spend visibility and analysis plan is not in place, new opportunities will not be identified and savings will decrease rapidly, in line with the curve in Figure 1.

![“Classic” Analysis Value Curve](image)

*Figure 1 The "Classic" Spend Analysis Value Curve*
However, if a strategic spend visibility and analysis plan is put in place, analysis will go beyond simple top-N reports and traditional procurement categories. It will also focus on identifying overcharges and overpayments that can only be detected by going beyond classic Accounts Payable (AP) data and analyzing invoices. From invoices, fraud detection, regulatory compliance analysis, identification of “sacred cow” spending in marketing and legal, product and part rationalization, and risk reduction opportunities become evident. Each of these initiatives, which will collectively make use of all of the organizational data sources that impact spending (and not just AP data), will identify new savings opportunities and initiate a new savings cycle. As a result, the organization will see year over year over year savings, in line with the curve in Figure 2.

![Figure 2 The Modern Spend Analysis Value Curve](image-url)
A Need for Strategy

As a result of the declining value curve in tactical spend visibility and analysis initiatives, a strategic approach is needed to ensure that savings opportunities will be maximized year after year. This strategic approach starts with an implementation plan that insures the proper foundations are in place for repeatable and realizable savings opportunities.

Basic Process

The basic mechanics of spend analysis only takes five or six easy steps, but a true strategic spend visibility initiative is a ten-step commitment. Each step is required to properly frame the project and build a foundation for long term savings success. Skipping the introductory steps and going straight to the data collection phase reduces spend analysis to a tactical initiative, and this is why many adopters of spend visibility and analysis systems see their savings opportunities drop sharply in the second or third year. Only those who execute each and every step will see the full set of savings opportunities available to them. This section will walk through each of these steps and provide examples where relevant.

01 Get Executive Support
02 Understand Existing Capabilities and Gaps
03 Identify the Data Sources
04 Identify the Raw Data
05 Define the Schema
06 Define the Starting Rules for Classification and Mapping
07 Avoid the Common Tactical Traps
08 Define the Cubes
09 Analyze, Assess, Report, Decide
10 Refresh, Repeat

01 Get Executive Support

This is one of the most crucial steps to the implementation of a successful spend visibility and analysis system and the one that’s most often overlooked. Often, an organization that is just beginning its spend visibility journey will assume that spend analysis is solely a procurement exercise and that as long as procurement personnel have access to AP data, they don’t need any organizational buy-in.

The reality is that a successful spend visibility effort requires spend and spend related data from across the organization as well as the support of the business units that own that data. This is true not just because the support of the business units might be required for the extraction of the data, but also because the business units might fight the initial adoption and implementation of the solution if they are afraid that they will be faulted for spending inefficiencies the solution might expose.
They may also believe that the implementation of the system is the first step to reducing their spend control. Even if neither of these things is true, executive support will be required to overcome business unit resistance. That’s why acquiring executive support is the first step in any spend visibility undertaking, because it will prevent otherwise intractable problems from arising.

What are you asking from these executives? Three simple things:
1. Provide feedback, until they trust the data
2. Use the data to generate value in their organization
3. Give suggestions on how to improve or expand the data to better meet their needs over time

Over time, you will go back to these same executives and show them the value that the system is delivering or supporting. These executives are the people who will be generating value from the data. As these executives see the results, they will continue to push for adoption and push to have the organization take advantage of the data. Having perfect data in a box does nothing for your organization. To get value, people must use the data. Over time, spend analysis must become part of “the way we do things”.

02 Understand Existing Capabilities and Gaps

In order for a spend visibility initiative to be successful, it has to exceed current organizational capabilities for spend visibility and analysis and address the gaps that are preventing the organization from achieving its savings opportunities.

Start by identifying the spend related systems that are in place, the spending data that is tracked, the reports that are produced and how this data is used or could be used today. What information can be extracted, what intelligence can be obtained, and what initiatives can be supported? For example, if the organization uses a single Oracle system as its ERP (Enterprise Resource Planning) and it has Oracle Financials, then the organization will have its accounts payable data centralized and will be able to run simple top-n reports, e.g. top 20 vendors for a GL (General Ledger) code or cost center. These reports sound like spend analysis, but they have significant limits. Vendors will show up as they are in the source system – with no vendor grouping (described in later sections). There will be limited commodity coding of data. GL codes do not fully line up to commodities. If the ordering system allows entry of commodity codes, codes may exist, but they will only be as good as the original data entry.
In a company with a single ERP system, if a fair number of purchases are made on P-Cards, chances are there will be no visibility into the P-Card payments beyond the amount paid to (say) American Express. If P-Cards account for 20% of spend, visibility into P-Card data would be a huge savings opportunity, especially if such spending would allow additional leverage with the existing supply base. This would be the case if 30% of office supplies or 25% of laptops were being bought off-contract on P-Card.

However, if the organization did not have Oracle iProcurement or the iSupplier portal, and instead used a back-office solution, then invoice data would not be readily available for analysis. Considering that invoice data is required to detect overbilling, off-contract purchases, and fraudulent billings, this would be a gap that any spend visibility system would need to address in order to access additional savings opportunities.

While the specific capabilities required to address the gaps in the organization’s systems and data will be dependent upon the current systems the organization has in place, the following is a brief set of capabilities, by type, that should be present in any spend analysis system under consideration. A system with these capabilities will likely be able to address whatever gaps are present in spend analysis capabilities in an average organization.

**DATA LOAD:**
- Ability to load data from multiple data sources in different formats. Must not require a fixed format.
- Ability to cope with missing or incomplete data *without ever dropping transactions.*
- Ability to load and link related data like vendor master or organizational hierarchy via fixed or formulaic keys (on either side of the join).
- Ability to join data across multiple relations. (i.e. link to link to link)
- Ability to transform data from multiple sources into a single format, and to store the transformation as a repeatable script, without learning a scripting language or a complex ETL (Extract-Transform-Load) tool.
- Ability to build arbitrary dimensions without fixed names or organizations – e.g. multiple organizational views or customer-designated geographic organization.

**DATA TRANSFORMATION:**
- Ability to group vendors – both manually and by leveraging external data – in real time.
- Ability to create and modify hierarchical structures in real time (Figure 3)
Figure 3 Changing Hierarchies

Telecom under IT Operations

Telecom under G&A
- Ability to map spending to commodity based on hierarchically-applied business rules – online with the cube, in real time, with full visibility into all dimensions, and with instant feedback on rule effects.
- Ability to build rule structures on any dimension, change keys (Figure 4), and control rules application where one mapped dimension depends on another.
- Ability to automatically track and update the mapping rules when the original target hierarchy structures are altered.
- Ability to do calculations across the data at a transactional, nodal or aggregate basis, and create arbitrary measures (Figure 5).

**Figure 5 Creating Arbitrary Measures**

**REPORTING AND OUTPUT:**
- Ability to report on all source data.
- Ability to extract all transactions and any results of the transformations.
- Ability to report on all the rules themselves, including the rules that affect a particular dimensional node.
- Ability to automatically populate and refresh existing Excel models with dimensional or multidimensional data from the spend cube.
- Ability to replicate existing Excel models into multi-page reports, booked on arbitrary dimensions.
- Ability to output transactional and/or dimensional data into Excel, CSV and other formats which can then be incorporated into further analysis and presentations.
- Ability to customize output quickly and easily without IT support (Figure 6). Some example reports are detailed in Figure 13.

**Figure 6 Building Advanced Reports**

**DATA REFRESH:**
- Ability to load incremental data and transform it as necessary, using previously-stored scripts or by creating new transforms.
- Ability to apply existing logic automatically including all rule operations to new data.
- Ability to ‘thin’ indexes, automatically updating just the entries that have changed in the index and maintaining all old items, such as orphan cost center not found in the current organizational hierarchy.
- Ability to collaborate quickly and easily with other users, sharing and merging hierarchy and mapping changes, as well as complete datasets.
- Ability to modify historical transformations of data (Figure 7 – in this example, both rearranging the field order and adding a system identifier to the codes).
Figure 7 Changing Data Formats

PERFORMANCE AND FLEXIBILITY:
- Fast filtering and searching capabilities. The response time must not get in the way of analysis. Be wary of systems where one has to “go get a cup of coffee while the machine thinks”.
- Ability to make changes to the input and output at any time.
- Multi-threaded operation.
- Automatic use of 64-bit architectures.
- Ad-hoc data analysis must be easy and fast. (Figure 8)

**Figure 8 Expected Value for Ad-hoc Data Analysis**

**USER CONTROL:**
- All of the above to be easily done without vendor or IT support. If every new report or dataset requires a work order for staff to do the work, the work won’t get done.

**Understand your users and potential users**

In successful companies, spend analysis is done broadly throughout the organization. People understand the need for data, have the skills to work with it and are looking to take advantage of it. In every company we look at, there are many people who work with procurement data. Even if a company doesn’t have a spend analysis system, there are people exporting blocks of data and crunching spreadsheets.
- **Who is using the data today?** Who are the people in procurement and the business units who are using the data today? These are the people who will be crucial in understanding your data and they will be your first users and advocates for the system. Putting in a spend analysis system that “makes analysis easy” may feel threatening to them if their job is to be “the data person”. It’s a common fear. In reality, the more people who use the data, the more value people who understand the data can deliver to their organization. But change can be frightening, so you need to work to prevent pushback from these people.

- **Who should be using this data?** Spend data is broadly useful. Sourcing team members. Sourcing management. Procurement operations. Procurement systems. Accounts Payable. Department heads. Project managers with budget responsibility. Senior executives. The list goes on. Many of these people are not going to focus on the details of the data, they will just want the “right” reports. Others will want to go into the data in detail. You should understand your initial user community.

- **Do they understand what this data can mean for them?** Spend analysis is done to deliver value. You should understand how the system will deliver value to your initial users. This will help you communicate the value, focus your training and ensure you aren’t missing key elements in the system. If a key value is to track bypass spending, it will be important for the system to have a measure of “managed vendors” vs. “non-managed vendors”.

- **Do they have the skills to work with the data?** You should understand how much training your initial user community will need. Once you have your spend analysis system in place, will they be “lions going after a hunk of red meat” or will you need to walk them through the process?

- **Do they have the time to work with the data?** Users will sometimes say that the data is useful but they don’t have time to work with it. Sometimes this is fear talking, so training and communication will reduce this pushback. A good spend analysis system will reduce the “time to useful information”. For example, clear Executive expectations that every sourcing team must present data and facts will help people to understand the need and the value of data, so they will make time for the system.

Understanding your user community will provide critical insight into your deployment approach and timing. A typical deployment path is to start with a small team and gradually expand the user community over time. Other organizations insist that all sourcing staff use the data and the tool immediately. What will be successful in your organization will depend on many factors, so you need to be aware of these factors.
03 Identify the Data Sources

Once the gaps have been identified, it’s time to identify all of the data sources that contain spend data and spend-related data that can be used in strategic analysis efforts. Specifically, it is important to identify what are the data sources, where are they located, who is responsible for them, when is updated data available for analysis, how is the data accessed, and why the data sources are important.

The identified data sources should contain enough data to classify the majority of organizational spending by category and supplier, to dive into usages by division, department, and geographic unit, and to permit N-way matching between contract data, purchase orders, invoices, goods receipts, and payments.

Figure 9 outlines some common organizational data sources that should be tracked down, and included, if present, as part of a successful spend analysis project.

<table>
<thead>
<tr>
<th>Type of Data</th>
<th>Typical Core Data</th>
<th>Additional Potential Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spending Data</td>
<td>AP Data</td>
<td>• Procurement Card (P-Card)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Credit Card</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Wire Transfers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ACH / EFT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Employee Reimbursements</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Payroll deductions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Contra Accounts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• “cash tickets” (retailers and banks)</td>
</tr>
<tr>
<td>Purchase Order</td>
<td>PO Data</td>
<td>• Legal Matters</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Third party authorizations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• (e.g. facilities management)</td>
</tr>
<tr>
<td>Contracts</td>
<td>Contract Systems</td>
<td>• File boxes of paper files (Scan/OCR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• (Microsoft) Access databases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Vendor databases</td>
</tr>
<tr>
<td>Projects</td>
<td>Project Tracking System</td>
<td>• Spreadsheets with project data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Management reports of projects</td>
</tr>
</tbody>
</table>

Figure 9 Organizational Data Sources for Spend Analysis

You may choose not to load all the data from all the sources into the system, or you may load only parts of individual systems. You may have organizational pushback. You may have resource limits. You may start with a single geography or just the AP data. You want to load enough data to be meaningful and to allow you to deliver value. Don’t wait a year for “perfect data” when you can get good data next week. Getting a simple extract from the most AP systems is quick and easy. A good spend analysis system should allow you to add more data later. What is enough data? Enough to deliver some value.
04 Identify the Raw Data

After the data sources are identified, the next step is to identify the data that is available for extraction. For example, most AP systems will be able to provide supplier, GL code, cost center, currency, payment amount, and payment date. Depending on the particular AP system used and how it is configured, additional fields might include business unit, country, geographic location, purchaser, authorizing manager, description, payment method, clearing date, purchase order number, legal entity, and asset ID. It’s important to identify all of the data that may be available before any attempts to extract it are made, because the more relevant the data is that is available, the more successful the spend visibility initiative will be. Sometimes the only way to properly classify transactions which are light on detail is to use a triangulation approach that matches the invoice to the purchase order and goods receipt to identify and extract all of the data corresponding to a transaction.

Figure 10 shows how AP data might be available in an organization that employs both SAP and PeopleSoft systems in different divisions.

Figure 11 shows what kind of invoice data might be available to an organization that employs an Oracle ERP solution to track invoice data.
### Example Data from AP Source: SAP

<table>
<thead>
<tr>
<th>AP FIELD NAME</th>
<th>SAMPLE DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2010</td>
</tr>
<tr>
<td>Period</td>
<td>08</td>
</tr>
<tr>
<td>Co Code</td>
<td>001028</td>
</tr>
<tr>
<td>Doc #</td>
<td>0017004842</td>
</tr>
<tr>
<td>Doc Type</td>
<td>VI</td>
</tr>
<tr>
<td>Doc Description</td>
<td>Vendor Invoice</td>
</tr>
<tr>
<td>Doc Status</td>
<td></td>
</tr>
<tr>
<td>Posting Date</td>
<td>20100826</td>
</tr>
<tr>
<td>Doc Date</td>
<td>20100216</td>
</tr>
<tr>
<td>Entered Date</td>
<td>20100828</td>
</tr>
<tr>
<td>Baseline Date</td>
<td>20100216</td>
</tr>
<tr>
<td>Vendor #</td>
<td>0000107607</td>
</tr>
<tr>
<td>Vendor Name</td>
<td>Boston Consulting Group</td>
</tr>
<tr>
<td>Text Description</td>
<td>Consulting Spend for Project Zenith</td>
</tr>
<tr>
<td>Cost Center</td>
<td>456213</td>
</tr>
<tr>
<td>Profit Center</td>
<td>006</td>
</tr>
<tr>
<td>G/L Acct #</td>
<td>0000211099</td>
</tr>
<tr>
<td>Account Name</td>
<td>Business Consulting</td>
</tr>
<tr>
<td>$ Amount</td>
<td>65000.00</td>
</tr>
<tr>
<td>Currency Code</td>
<td>EUR</td>
</tr>
<tr>
<td>PO Doc #</td>
<td></td>
</tr>
<tr>
<td>PO Doc Item</td>
<td></td>
</tr>
<tr>
<td>Int Order #</td>
<td>STRAT34</td>
</tr>
<tr>
<td>Invoice Number</td>
<td>BCG45345</td>
</tr>
</tbody>
</table>

### Example Data from AP Source: PeopleSoft

<table>
<thead>
<tr>
<th>AP FIELD NAME</th>
<th>SAMPLE DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>BUSINESS_UNIT_AP</td>
<td>20020</td>
</tr>
<tr>
<td>VOUCHER_ID</td>
<td>89459978</td>
</tr>
<tr>
<td>VENDOR_ID</td>
<td>4654852</td>
</tr>
<tr>
<td>REMIT_ADDR_SEQ_NUM</td>
<td>2</td>
</tr>
<tr>
<td>VENDOR_NAME1</td>
<td>VERTEX BUSINESS SERVICES</td>
</tr>
<tr>
<td>ADDRESS1</td>
<td>PO BOX 850001</td>
</tr>
<tr>
<td>ADDRESS2</td>
<td></td>
</tr>
<tr>
<td>CITY</td>
<td>ORLANDO</td>
</tr>
<tr>
<td>STATE</td>
<td>FL</td>
</tr>
<tr>
<td>POSTAL</td>
<td>32885-0001</td>
</tr>
<tr>
<td>INVOICE_ID</td>
<td>REIMBVERTEX108564</td>
</tr>
<tr>
<td>INVOICE_DT</td>
<td>10/29/2009</td>
</tr>
<tr>
<td>GROSS_AMT</td>
<td>150040.69</td>
</tr>
<tr>
<td>PYMNT_STATUS</td>
<td>P</td>
</tr>
<tr>
<td>BUSINESS_UNIT_GL</td>
<td>45682</td>
</tr>
<tr>
<td>Legal BU</td>
<td>Samosome</td>
</tr>
<tr>
<td>OPERATING_UNIT</td>
<td>8546</td>
</tr>
<tr>
<td>CHARTFIELD1</td>
<td>USAHI802</td>
</tr>
<tr>
<td>ACCOUNT</td>
<td>59999</td>
</tr>
<tr>
<td>ACCOUNT_DESCR</td>
<td>Miscellaneous Benefits Paid</td>
</tr>
<tr>
<td>DEPTID</td>
<td>310220</td>
</tr>
<tr>
<td>PRODUCT</td>
<td>250022</td>
</tr>
<tr>
<td>CHARTFIELD2</td>
<td>8100000</td>
</tr>
<tr>
<td>CHARTFIELD3</td>
<td></td>
</tr>
<tr>
<td>PROJECT_ID</td>
<td></td>
</tr>
<tr>
<td>MONETARY_AMOUNT</td>
<td>18892.3</td>
</tr>
<tr>
<td>APPROVER_NAME</td>
<td>SMITH, M,EID:43534</td>
</tr>
<tr>
<td>DESCRIPT</td>
<td>TOTAL AMOUNT DUE</td>
</tr>
<tr>
<td>PO_ID</td>
<td>1423</td>
</tr>
<tr>
<td>PYMNT_DT</td>
<td>1/7/2010</td>
</tr>
<tr>
<td>PYMNT_ID_REF</td>
<td>31176766</td>
</tr>
<tr>
<td>SCHEDULED_PAY_DT</td>
<td>1/6/2010</td>
</tr>
<tr>
<td>PYMNT_METHOD</td>
<td>CHK</td>
</tr>
<tr>
<td>PYMNT_HOLD</td>
<td>N</td>
</tr>
<tr>
<td>OPRID</td>
<td>PROCSCHD</td>
</tr>
<tr>
<td>ORIGIN</td>
<td>CON</td>
</tr>
<tr>
<td>ASSET_ID</td>
<td></td>
</tr>
<tr>
<td>SALETX_AMT</td>
<td>0</td>
</tr>
<tr>
<td>FREIGHT_AMT</td>
<td>0</td>
</tr>
<tr>
<td>FISCAL_YEAR</td>
<td>2010</td>
</tr>
<tr>
<td>ACCOUNTING_PERIOD</td>
<td>1</td>
</tr>
<tr>
<td>ACCOUNTING_DT</td>
<td>1/6/2010</td>
</tr>
<tr>
<td>COMMODITY_DESCR</td>
<td>Real Estate Services</td>
</tr>
<tr>
<td>SHIP TO LOC</td>
<td>MEMPH</td>
</tr>
</tbody>
</table>

---

**Figure 10 AP Record Examples**
<table>
<thead>
<tr>
<th>Invoice Field Name</th>
<th>Sample Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sold To</td>
<td>10240592</td>
</tr>
<tr>
<td>Order Number</td>
<td>310310738</td>
</tr>
<tr>
<td>Order Item</td>
<td>10</td>
</tr>
<tr>
<td>Order Item Date</td>
<td>12/10/2010</td>
</tr>
<tr>
<td>Invoice Number</td>
<td>900378211</td>
</tr>
<tr>
<td>Invoice Item</td>
<td>10</td>
</tr>
<tr>
<td>Invoice Date</td>
<td>12/12/2010</td>
</tr>
<tr>
<td>Ship-To</td>
<td>10240592</td>
</tr>
<tr>
<td>Ship-To Name</td>
<td>COMPANY</td>
</tr>
<tr>
<td>Name 2</td>
<td></td>
</tr>
<tr>
<td>Ship-To C/O</td>
<td>CHARLES WHITE</td>
</tr>
<tr>
<td>Street</td>
<td>512 Main St</td>
</tr>
<tr>
<td>City</td>
<td>Arlington</td>
</tr>
<tr>
<td>Region</td>
<td>MA</td>
</tr>
<tr>
<td>Postal Code</td>
<td>01712</td>
</tr>
<tr>
<td>Material Description</td>
<td>Software - Software Service/Support</td>
</tr>
<tr>
<td>Material</td>
<td>PRD07630010</td>
</tr>
<tr>
<td>Manufacturer Part Number</td>
<td>BES CAL 1U 4.1V F 3.6/2.2</td>
</tr>
<tr>
<td>Manufacturer</td>
<td>BLACKBERRY</td>
</tr>
<tr>
<td>Serial Number</td>
<td>#</td>
</tr>
<tr>
<td>Customer PO Number</td>
<td>DK451032</td>
</tr>
<tr>
<td>Quantity</td>
<td>1 EA</td>
</tr>
<tr>
<td>Unit Price</td>
<td>$ 90.00 /EA</td>
</tr>
</tbody>
</table>

Figure 11 Invoice Example
05 Define the Schema

Once all of the raw data has been identified, the next step is to create a data format that will allow the data elements to be integrated into a single master cube for spend analysis. This will allow for the creation of the master files that will define unique suppliers, categories, products, and departments, among other entities.

Why Do We Need A “Master” Cube? Remember the Basic Question

The ultimate goal of spend visibility is to answer who is buying what from whom, when, in what quantity, where they are shipping it, and how are they paying for it. The only question that isn’t being asked is why, since the answer to that question should always be “because it is needed”. The data format that is defined should be able to fully answer this question.

Who: Cost center (organizational roll up)
What: GL Account, Account Name, Text Description
Where: Ship To Location
When: Invoice Date
How: PO ID, Payment Method
From whom: Vendor

Figure 12 below illustrates the basic who buys what from whom analysis.
Without a master schema that relates all of the disparate spend data throughout the organization, even a simple “who buys what from whom” query might be beyond the ability of an organization to grasp.

At the end of the day, the organization should be able to produce a clustered crosstab report, like the one developed by the original Mitchell Madison Group in the mid 1990’s (captured in Figure 13), that will allow it to answer all of these questions.
The Clusters Crosstab Report shows different slices of data for a single point in the dataset. The example in Figure 13 shows the “One Page Per Commodity” view for data. Each page is for a single commodity. For each commodity, you see some summary data (e.g. total spend, total transactions, number of vendors, number of GL codes, number of organizational units, etc.), a list of top vendors, a list of the top GL codes, a list of organizational units, a list of the top cost centers and a time series of the data. For each list, you see the names and spend for each name. This gives you a quick, single page overview of a commodity with everything you need to know in one place.

Creating a series of these reports, shown in Figure 13b, gives a quick overview across multiple commodities. You can filter on a single business unit and create the reports for this business unit. A business leader can flip through a stack of these reports and see their overall spending patterns.
Figure 13b– A series of The Clustered Crosstab Report

This report can be done in many different ways. Figure 13c shows an alternative display in which the panels provide the context for the spending in the report. In this example, for each vendor, the report shows the spending for this single commodity with a second column showing the overall spending for this vendor. For vendors with partial spending, the reader quickly understands that there is more spending with that vendor that exists in other commodities. This gives some “peripheral vision” into how much this partial page represents of the total vendor spending (or the single GL or Cost Center).
The table in Figure 14 illustrates how relevant AP data might be merged from SAP (System 1) and PeopleSoft (System 2) systems into a common AP data format. Notice in the combined data view how key code fields have been made unique by pre-pending the system ID to the codes.
### System 1

<table>
<thead>
<tr>
<th>Field</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>VENDOR #</td>
<td>107607</td>
</tr>
<tr>
<td>VENDOR NAME</td>
<td>BOSTON CONSULTING GROUP</td>
</tr>
<tr>
<td>INVOICE_ID</td>
<td>BCG45.845</td>
</tr>
<tr>
<td>$ AMOUNT</td>
<td>65000</td>
</tr>
<tr>
<td>CURR CODE</td>
<td>EUR</td>
</tr>
<tr>
<td>G/L ACCT #</td>
<td>211099</td>
</tr>
<tr>
<td>ACCOUNT NAME</td>
<td>BUSINESS CONSULTING</td>
</tr>
<tr>
<td>INVOICE_DT</td>
<td>20100826</td>
</tr>
<tr>
<td>TEXT DESCR</td>
<td>CONSULTING SPEND FOR PROJECT ZENITH</td>
</tr>
<tr>
<td>COST CENTER</td>
<td>456213</td>
</tr>
</tbody>
</table>

### System 2

<table>
<thead>
<tr>
<th>Field</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>VENDOR #</td>
<td>6664552</td>
</tr>
<tr>
<td>VENDOR NAME</td>
<td>VERTERX BUSINESS SERVICES</td>
</tr>
<tr>
<td>INVOICE_ID</td>
<td>REIMBVERTEX108564</td>
</tr>
<tr>
<td>$ AMOUNT</td>
<td>18892.3</td>
</tr>
<tr>
<td>G/L ACCT #</td>
<td>589999</td>
</tr>
<tr>
<td>ACCOUNT NAME</td>
<td>MISCELLANEOUS BENEFITS PAID</td>
</tr>
<tr>
<td>INVOICE_DT</td>
<td>40118</td>
</tr>
<tr>
<td>PYMENT_DT</td>
<td>40185</td>
</tr>
<tr>
<td>PO #</td>
<td>1423</td>
</tr>
<tr>
<td>TEXT DESCR</td>
<td>REAL ESTATE SERVICES</td>
</tr>
<tr>
<td>PYMENT_ID_REF</td>
<td>31176766</td>
</tr>
<tr>
<td>SCHEDULED_PAY_DT</td>
<td>40184</td>
</tr>
<tr>
<td>PYMENT_METHOD</td>
<td>CHK</td>
</tr>
<tr>
<td>PYMENT_HOLD</td>
<td>N</td>
</tr>
<tr>
<td>ORIGIN</td>
<td>CON</td>
</tr>
<tr>
<td>COST CENTER</td>
<td>310220</td>
</tr>
<tr>
<td>SHIP TO LOCATION</td>
<td>MEMPH</td>
</tr>
<tr>
<td>PAYMEN_METHOD</td>
<td>CHK</td>
</tr>
</tbody>
</table>

### Combined Data

<table>
<thead>
<tr>
<th>Field</th>
<th>Data</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>System</td>
<td>System 1</td>
<td>System 2</td>
</tr>
<tr>
<td>VENDOR #</td>
<td>SYS1:107607</td>
<td>6664552</td>
</tr>
<tr>
<td>VENDOR NAME</td>
<td>BOSTON CONSULTING GROUP</td>
<td>VERTERX BUSINESS SERVICES</td>
</tr>
<tr>
<td>INVOICE_ID</td>
<td>BCG45.845</td>
<td>REIMBVERTEX108564</td>
</tr>
<tr>
<td>$ AMOUNT</td>
<td>65000</td>
<td>18892.3</td>
</tr>
<tr>
<td>CURR CODE</td>
<td>EUR</td>
<td>EUR</td>
</tr>
<tr>
<td>G/L ACCT #</td>
<td>SYS1:211099</td>
<td>SYS2:589999</td>
</tr>
<tr>
<td>ACCOUNT NAME</td>
<td>BUSINESS CONSULTING</td>
<td>MISCELLANEOUS BENEFITS PAID</td>
</tr>
<tr>
<td>INVOICE_DT</td>
<td>20100826</td>
<td>40118</td>
</tr>
<tr>
<td>PYMENT_DT</td>
<td>40185</td>
<td>40185</td>
</tr>
<tr>
<td>PO #</td>
<td>1423</td>
<td>1423</td>
</tr>
<tr>
<td>TEXT DESCR</td>
<td>CONSULTING SPEND FOR PROJECT ZENITH</td>
<td>REAL ESTATE SERVICES</td>
</tr>
<tr>
<td>PYMENT_ID_REF</td>
<td>31176766</td>
<td>31176766</td>
</tr>
<tr>
<td>SCHEDULED_PAY_DT</td>
<td>40184</td>
<td></td>
</tr>
<tr>
<td>PYMENT_METHOD</td>
<td>CHK</td>
<td>CHK</td>
</tr>
<tr>
<td>PYMENT_HOLD</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>ORIGIN</td>
<td>CON</td>
<td>CON</td>
</tr>
<tr>
<td>COST CENTER</td>
<td>SYS1:456213</td>
<td>SYS2:310220</td>
</tr>
<tr>
<td>SHIP TO LOCATION</td>
<td>MEMPH</td>
<td></td>
</tr>
<tr>
<td>PAYMEN_METHOD</td>
<td>CHK</td>
<td>CHK</td>
</tr>
</tbody>
</table>

**Figure 14** An AP Record Merge Example
Notice how the merged data format provided in Figure 14 answers the who – what – where – how – from whom question.

**Merging Similar Datasets**

The most common scenario for merging data from two or more fundamentally similar data sources into a common data format for spend analysis is when multiple AP systems are present within the same company. A five-step process is required.

1. **Identify the Unique Fields**
   The first step is to identify and list all of the unique fields that will form the merged record. This is a manual effort, best performed with a simple spreadsheet (listing all the fields) or even with pencil and paper.

2. **Disambiguate Key Fields**
   Numeric fields in AP systems, such as Vendor ID, will mean different things in different AP systems. Vendor 137 might mean IBM in Subsidiary A's Oracle system, but something entirely different in Subsidiary B's JD Edwards system. This situation is common in companies that have undergone acquisitions. Obviously, these key values must be made unique; and the easiest way to do this is to prepend a notion of system ID to the field.

3. **Define and Align the Field Formats**
   If field format matters, such as with dates, then multiple formats should be mapped to a consistent format.

4. **Map Existing Fields**
   After applying disambiguation rules and format conversions, align common fields with each other, inserting additional empty (or fixed value) fields where necessary.

5. **Define Constants Where Required**
   Some required values may not exist in a source file but need to have a value in the output. For example, if all transactions in the AP system are in USD, then the currency field, if blank, can be defaulted to USD. In addition, an indication as to the source of the data should be included, so that the data attributable to a particular system can be identified later.

A modern spend analysis system will be able to define and automate steps 2 through 5, so that they are trivially repeatable when data need to be refreshed.
Merging Disparate Datasets

When disparate datasets need to be analyzed together – and when federation between those disparate datasets is not an option provided by the spend analysis system – then there is no choice but to transform the disparate transactions from both sources into a common uber-transaction. A good example of this is when an analyst wants to explore both AP and PO data in the same cube. The AP transactions contain actual payments to vendors, which include vendor name, GL information, cost center, date, and amount paid, as well as PO number if payments were made against an outstanding PO. The PO transactions contain some similar information, but the PO amount is merely an amount authorized, not an actual amount paid. These are really two very different datasets.

To combine two disparate datasets with one or more common dimensions, the process is as follows:

1. **Create a common record format.**
   Define a combined transaction record format that contains all unique fields from both data sources. If there are identical fields, only one copy of the field is required. Make sure that a new SourceID field is included in the common record, so that later on the dataset can be filtered (drilled) by source. This is a manual effort easily done with two columns on a spreadsheet.

2. **Create transforms for each data source**
   Define transforms that convert records from all data sources into the common record format.

3. **Load all of the transformed sources into a new dataset.**

A modern spend analysis system will be able to define and automate step 2, as indicated above.

An example of this is depicted in Figure 15.
### AP Data (transformed)

<table>
<thead>
<tr>
<th>Field</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>VENDOR #</td>
<td>107607</td>
</tr>
<tr>
<td>VENDOR NAME</td>
<td>BOSTON CONSULTING GROUP</td>
</tr>
<tr>
<td>INVOICE_ID</td>
<td>BC645345</td>
</tr>
<tr>
<td>AMOUNT</td>
<td>65000</td>
</tr>
<tr>
<td>CURR CODE</td>
<td>USD</td>
</tr>
<tr>
<td>INVOICE_DT</td>
<td>8/26/2010</td>
</tr>
<tr>
<td>COST CENTER</td>
<td>456213</td>
</tr>
<tr>
<td>G/L ACCT #</td>
<td>211099</td>
</tr>
<tr>
<td>ACCOUNT NAME</td>
<td>BUSINESS CONSULTING</td>
</tr>
<tr>
<td>PO Number</td>
<td>PO534023</td>
</tr>
<tr>
<td>PO Date</td>
<td></td>
</tr>
<tr>
<td>PO AMOUNT</td>
<td></td>
</tr>
<tr>
<td>PO CURR</td>
<td></td>
</tr>
<tr>
<td>TEXT DESCR</td>
<td></td>
</tr>
<tr>
<td>Record Type</td>
<td>AP</td>
</tr>
</tbody>
</table>

### PO Data (transformed)

<table>
<thead>
<tr>
<th>Field</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>VENDOR #</td>
<td>107607</td>
</tr>
<tr>
<td>VENDOR NAME</td>
<td>BOSTON CONSULTING GROUP</td>
</tr>
<tr>
<td>INVOICE_ID</td>
<td>BC645345</td>
</tr>
<tr>
<td>AMOUNT</td>
<td>325000</td>
</tr>
<tr>
<td>CURR CODE</td>
<td>USD</td>
</tr>
<tr>
<td>INVOICE_DT</td>
<td>6/4/2010</td>
</tr>
<tr>
<td>COST CENTER</td>
<td></td>
</tr>
<tr>
<td>G/L ACCT #</td>
<td></td>
</tr>
<tr>
<td>ACCOUNT NAME</td>
<td></td>
</tr>
<tr>
<td>PO Number</td>
<td>PO534023</td>
</tr>
<tr>
<td>PO Date</td>
<td>8/26/2010</td>
</tr>
<tr>
<td>PO AMOUNT</td>
<td>325000</td>
</tr>
<tr>
<td>PO CURR</td>
<td></td>
</tr>
<tr>
<td>TEXT DESCR</td>
<td>CONSULTING SPEND FOR PROJECT ZENITH</td>
</tr>
<tr>
<td>Record type</td>
<td>PO</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Field</th>
<th>AP Data (transformed)</th>
<th>PO Data (transformed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VENDOR #</td>
<td>107607</td>
<td>107607</td>
</tr>
<tr>
<td>VENDOR NAME</td>
<td>BOSTON CONSULTING GROUP</td>
<td>BOSTON CONSULTING GROUP</td>
</tr>
<tr>
<td>INVOICE_ID</td>
<td>BC645345</td>
<td>BC645345</td>
</tr>
<tr>
<td>AMOUNT</td>
<td>65000</td>
<td>65000</td>
</tr>
<tr>
<td>CURR CODE</td>
<td>USD</td>
<td>USD</td>
</tr>
<tr>
<td>INVOICE_DT</td>
<td>8/26/2010</td>
<td>8/26/2010</td>
</tr>
<tr>
<td>COST CENTER</td>
<td>456213</td>
<td>456213</td>
</tr>
<tr>
<td>G/L ACCT #</td>
<td>211099</td>
<td>211099</td>
</tr>
<tr>
<td>ACCOUNT NAME</td>
<td>BUSINESS CONSULTING</td>
<td>BUSINESS CONSULTING</td>
</tr>
<tr>
<td>TEXT DESCR</td>
<td></td>
<td>CONSULTING SPEND FOR PROJECT ZENITH</td>
</tr>
<tr>
<td>PO Date</td>
<td></td>
<td>8/26/2010</td>
</tr>
<tr>
<td>PO AMOUNT</td>
<td></td>
<td>325000</td>
</tr>
<tr>
<td>Record Type</td>
<td>AP Data</td>
<td>PO Data</td>
</tr>
</tbody>
</table>

**Figure 15 An AP and PO Record Merge Example**
Loading these records into a dataset results in two measure columns, for PO and for AP. Filtering on a single PO will show both the AP record(s) and the PO record associated with that PO number. Filtering on the "PO" record type shows just PO records; filtering on "AP" shows just AP records. Filtering on common dimensions, such as Vendor# results in a display of all AP and PO records associated with that Vendor.

**Special Case Merging**

The AP and PO example above is a simple one, so it is susceptible to merge approaches that work "sometimes" and in "some cases." Beware of "easy" solutions, because easy solutions create "not so easy" problems during analysis. For example, the PO data could be loaded as an index linked to AP (meaning that the PO is linked to every transaction that references it). PO AMOUNT is then (trivially) derived from the linked index, and there is no need to create a merged transaction, or to format each transaction stream, and so on.

This strategy works at a superficial level, but it will create duplicate PO "reality" for each AP transaction that bills against the PO, because the relationship of AP to PO is many-to-many. This means, for example, that the "PO AMOUNT" column may roll up to more than the actual PO amount, because multiple AP records that refer to the same PO will report the same amount, duplicating that amount (potentially) one or more times. Thus, rolled-up PO AMOUNTs in this cube will be unusable and untrustworthy.

Here is a simple example where there are two invoices for a total of $300 drawn against a PO with a total value of $500. If the system does a simple join of the records, each record in the AP system will get the PO Amount attached to it. When this new column of data is added up, the total for PO Amount will be $1,000, clearly the wrong amount.
You may want to live with this issue, by telling consumers of the dataset that "You have to ignore any rollups of PO AMOUNT" or some other story, but built-in errors in the cube will mean, inevitably, that someone will produce the wrong number at some point. Wrong numbers are dangerous, because they can easily cause serious consequences for you, the cube, and your entire spend management project.

There is a way to fix this, but it is complex. If the spend analysis system has a general-purpose Computed Measure facility, the problem can be addressed by creating a new COMPUTED PO AMOUNT variable (let's call it "CPAMT"). CPAMT is computed at the transaction level (using the Computed Measure facility) by dividing the PO AMOUNT of each transaction by the transaction count associated with the PO. The resulting (potentially fractional) CPAMT totals should sum up properly.

This is easier said than done, however, because the transaction count associated with the PO at the current filter position is not known until after we compute the current filter position. This is a "chicken-and-egg" problem that is common in data analysis. There is a solution to this, too. If the spend analysis system has a Dynamic Reference Filter facility, in other words a way to pre-compute the current filter position and store it away for reference, the chicken-and-egg problem can be solved by deriving a new dimension on PO number called PO_DIM. Its nodes in the Dynamic Reference Filter, before the actual rollup, will contain the count of the transactions that we need. The transactional Computed Measure formula for CPAMT then becomes:

\[
CPAMT = \frac{\text{PO AMOUNT}}{\text{lookup(DynamicReferenceFilter_1.PO_DIM, [PO number]).Count}}
\]

The formula states "Let CPAMT be the PO AMOUNT divided by the count of transactions stored in the just-previous-rolled-up-in-Dynamic-Reference-Filter-1's PO_DIM dimension, at the PO number referenced by this transaction."
Even this solution, however, is not perfect. For example, if two different Organizations had ordered from the same PO, the CPAMT measure would be the total amount for the PO, not the amount of the PO that this organization was responsible for. Displaying the total PO amount might or might not be what the organization is used to seeing.

You will need to apply the merge solution that works best based on your analysis needs, your source data, and exposes you to the least amount of external criticism.

**Address Critical Issues**

When merging data sources, a number of issues may arise when attempting to map source data. Here are some critical issues that need to be addressed.

**Currency**

It will often be the case that currency conversions will be required to merge data from multiple data sources. Most companies with a significant amount of spend in multiple currencies will need to track both the local currency and the reporting currency. In addition, some companies will require reporting in multiple currencies. For example, a multi-national might have US managers who need to see all spending in US Dollars, while their German counterparts might need to see spending in Euros. The system not only needs to accommodate this, but also needs to maintain the original currency and amount to allow users to explore the actual spending that occurred.

The currency conversion rates should be obtained from the organization’s finance group. The finance group should also inform the analysis team if the rates are reset for reporting on a daily, monthly or quarterly basis. Most organizations either provide the currency conversion in the source data or use a monthly average conversion rate.
Overlapping data

There are times where it is useful to pull together data from multiple, overlapping sources. In the case of PCard transactions, the summary payment to the PCard vendor may be contained in the AP transactions. The PCard detail may be more useful than this summary payment, because it contains information on the actual vendors used with each transaction. Blindly adding this data will cause the totals in the cube to be higher than the actual spending. This can be solved by adding a negative transaction to cancel out the PCard spending, removing the original PCard payments, or building a separate PCard dataset.

For some categories like facilities management, advertising agencies, technology purchases, etc., a vendor may manage or consolidate payments to a number of source vendors. You may want to know about the underlying spending. For example, if your advertising agency is managing a direct mail campaign, they may be using the same printers that you use for collateral materials. You may insist on this to ensure your trademarks get rendered appropriately. If the agency bills you for the campaign, you will not immediately see the total spending with the print vendor.

This problem of overlapping data can also occur in companies with multiple AP systems. It is critical to ensure that the data sources are well understood to avoid this type of problem.

Related Data – Indexes

Not all the data will be in the transaction file. There may be data about vendors in the vendor master. There may be organizational hierarchy information in a separate file. The transaction file may only contain GL Codes with the descriptions in a lookup table.

Technically, linking this extra data into the transaction file involves indexing this data. If the organization currently does analysis with spreadsheets, it will need to do the equivalent of a vlookup() in the spend analysis tool. The linkage is done by looking up values in the transaction file and finding those rows in the index file.

For example, the transaction file could have a vendor ID. This ID would be used to look up the data in the index file. In Figure 14, in the first system, Vendor ID 107607 was the ID for Boston Consulting Group. In general, linking from a transaction to an index may require a complex function on transaction files, on the one side, and a different complex function on the index files, on the other side. "Perfect" keys often do not exist, so ensure that your spend analysis system has the ability to create arbitrary functions on both sides of the linkage.
Where there are multiple systems, the keys need to be made unique in both the transaction file and in the index file. This is done by adding a system identifier, in this case “SYS1” and “SYS2”. So:

SYS1:107607 = Boston Consulting Group
SYS2:107607 = Xerox

Index data can be of any sort of data. Beside the data sources mentioned above, there may be contract information relating to each vendor, details from the PO system or the number of FTE’s per cost center. Once data is indexed into the transaction data, it can be used for any type of analysis.

Be careful of indexes that are not many-to-one between transaction and index. If a transaction could refer to more than one index record, how will the system determine which index record to use? Also, transaction records that are appropriately many-to-one with respect to index might be fine for pulling in labels and so on, but not so fine when pulling in amounts. See "Special Case Merging" for an example of how much trouble you can get into in this case.

**Updating Indexes**

Over time, the index data will need to be updated. New vendors are added to the vendor master. New cost centers are added to the cost center table. The roll up of the cost centers into the organization view will vary over time. The organization managing the index data may only keep the current view of the index – discarding the old values. You need to maintain information on these old values.

The classic case is for cost center roll ups. When a cost center is no longer in use, it is dropped from the organizational roll up. Why would the accounting department maintain the roll up for a cost center you can no longer book expenses to? However, historical transactions will continue to exist, so for the spend cube, you will need to maintain these historical roll ups.

A good system will “thin” the index properly. Simply replacing the index with the new values will lose information. The process of thinning an index will update the next values in an index, but where no new information is added, the old data will be maintained.

**Timing**

Data draws from different systems may have different timing. You will need to confirm that transactions for the month of February represent the same range of dates. For example, one system might use the accounting date for transactions. Another system may use the check date. These will be close, but you will need to determine if adjustments need to be made.
Don’t Get Confused By Language

When gathering the spend data for a large organization, an analyst will often discover that there are certain “islands” of systems that serve specific spending needs. For example, the Legal department may have a system for managing their vendors. Many legal departments will say that they are different and need to manage their vendors in unique ways. Some of these points are valid, but once the analyst digs deeper, she will discover that a well-run Legal department uses many of the core purchasing documents and practices that are common to all departments. They will just call the items by different names, as summarized in Figure 17.

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Purchasing Name</th>
<th>Legal Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Document with controls that authorizes a project or a single block of spending with a vendor.</td>
<td>Purchase Order</td>
<td>Matter</td>
</tr>
<tr>
<td>System to manage the spending authorizations.</td>
<td>Purchase Order System</td>
<td>Matter Management System</td>
</tr>
<tr>
<td>A document which describes the general terms and conditions between the buyer and the seller.</td>
<td>Master Services Agreement</td>
<td>Arrangement Letter</td>
</tr>
<tr>
<td>Standard for electronic submission of invoices.</td>
<td>EDI (Electronic Data Interchange)</td>
<td>LEDES (Legal Electronic Data Exchange Standard)</td>
</tr>
</tbody>
</table>

**Figure 17 Different Names for the Same Purpose**

The analyst may also find specific systems for managing facilities leases, equipment leases, and vendor management systems for contingent labor, payroll, and so on. These specialized systems often meet business needs better than generic systems such as ERP systems, so they are not going away, despite the best efforts of some commentators in the spend analysis space to wish them away. A good spend analysis approach needs to be able to gather and analyze data from all these different systems.

**Check Your Control Totals**

Once the data is loaded into the system, a check should be done to confirm that the data is complete and consistent with the expected values.

This involves a comparison with known totals. Review the total dollar amounts involved and compare them to control totals from the original transaction source (e.g. AP system). Ensure that the amounts and transaction volumes are the same order of magnitude that is expected.
If the person creating the extract does not have access to an independent source of data with control totals, she can get good data either from the finance function / department or, if she works for a public company, from the annual report. If the organization spends $500 million per year from a well-defined set of vendors and the spend data set has only $400 million dollars of transactions from those vendors, then there is an issue.

It’s important to note that AP spend in the spend analysis system (cash expenditures) will be different from the accounting treatment of spend for the same time period, which will also adjust for depreciation, include payments outside the AP system, and contain data on business unit exclusions. The chart in Figure 18 shows how the spending in the annual report, derived from the accounting systems, might be compared to the spending in the spend analysis system.

![Figure 18 Mapping Annual Report Spend to Cash in the System](chart.png)

In Figure 18, which is from a manufacturer, the company is growing and cash spending on equipment is higher than the depreciation amount. So, starting from the base spend of 100, the depreciation amount, shown in the annual report, is replaced with the net incremental spend on new equipment. This is calculated as the change in the balance sheet for equipment (capital account at the end of the year) – (capital account at the start of the year). If there is no change in this value, the new purchases must have equaled the amount of equipment depreciation. This example shows a case where the organization is buying more equipment now that it has historically, so the cash outlay is greater than the depreciation amount. In addition, the spending outside the AP system is excluded, and then business units for which there is no AP data are excluded as well. The net result is that 90% of spend should be present and accounted for. If the spend analysis system total is 90%, you will have confidence that you have captured all of the intended spending.
In addition to generating a baseline for the spend in the spend analysis system with respect to the accounting systems (or summary data in the annual report), it is important to perform a few quality control checks on the data to make sure the imports worked appropriately. One important check is a time-series analysis. In this type of analysis, major dimensions are compared across time to see if data is missing or if there are major changes in spend patterns that need to be investigated. This confirms that all new data added is consistent with the original data loaded into the system.

The time-series analysis can be performed by GL code, by vendor, by source system, by country, or by any other dimension where spend is expected to be more-or-less steady over time. In Figure 19, where the time-series is created on the GL code, the cells in yellow represent data that could be missing or erroneous. Why are there no payments for “Construction in Process” for six months? Why are there no consulting fees, legal fees, or rent payments for the last three months? Why did the occupancy expense drop by 98% in August? While it may be the case that some of the cells should be blank – because the Construction Vendor switched to bi-monthly billing or AP accidentally paid a million dollar occupancy expense invoice twice in July and then had that money credited to their account in August – some of the cells are likely erroneous.

<table>
<thead>
<tr>
<th>GL Accounts</th>
<th>TD 01/02</th>
<th>TD 02/02</th>
<th>TD 03/02</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADVERTISING &amp; PROMOTIONS - 5612</td>
<td>424,589</td>
<td>4,840,266</td>
<td>1,177,203</td>
</tr>
<tr>
<td>Construction in Process - 5647</td>
<td>2,084</td>
<td>2,605,365</td>
<td>1,042,547</td>
</tr>
<tr>
<td>External Labor - 5687</td>
<td>1,126,613</td>
<td>9,949,952</td>
<td>859,411</td>
</tr>
<tr>
<td>SERVICE BUREAU EXPENSES - 56205</td>
<td>551,524</td>
<td>5,766,560</td>
<td>1,925,278</td>
</tr>
<tr>
<td>Legal - 5038</td>
<td>1,276,384</td>
<td>4,946,260</td>
<td>822,580</td>
</tr>
<tr>
<td>Legal FEES - 56135</td>
<td>637,676</td>
<td>1,030,855</td>
<td>935,105</td>
</tr>
<tr>
<td>Consulting FEES - 5648</td>
<td>498,649</td>
<td>7,042,915</td>
<td>567,752</td>
</tr>
<tr>
<td>PUBLIC RELATIONS - 56178</td>
<td>150,675</td>
<td>8,963,854</td>
<td>403,211</td>
</tr>
<tr>
<td>OCCUPANCY EXPENSE - 56149</td>
<td>69,422</td>
<td>1,214,108</td>
<td>1,247,149</td>
</tr>
<tr>
<td>Supplies - 56216</td>
<td>391,355</td>
<td>2,078,736</td>
<td>884,777</td>
</tr>
<tr>
<td>Health BENEFITS - 56200</td>
<td>1,211,768</td>
<td>54,731</td>
<td>1,235,832</td>
</tr>
<tr>
<td>Equipment REPAIRS/MAINTENANCE - 5680</td>
<td>365,764</td>
<td>1,411,808</td>
<td>660,862</td>
</tr>
<tr>
<td>Computer &amp; Equipment - 5644</td>
<td>618,418</td>
<td>1,145,424</td>
<td>554,065</td>
</tr>
<tr>
<td>FACILITIES - 5688</td>
<td>399,795</td>
<td>863,836</td>
<td>879,251</td>
</tr>
<tr>
<td>Checkbook Printing - 5639</td>
<td>686,453</td>
<td>571,578</td>
<td>755,139</td>
</tr>
<tr>
<td>Fixed Assets in Process - 5689</td>
<td>69,936</td>
<td>7,175,687</td>
<td>49,191</td>
</tr>
</tbody>
</table>

Figure 19 A Time Series Analysis

If there are spikes in the spending or in the number of transactions, this might be evidence that duplicate data has been loaded into the system. Once the spikes are identified, research needs to be done to confirm the changes. Sometimes there is a problem, sometimes there isn’t.
Year-over-year reports are very important when generating a spend baseline. Significant changes generally indicate potential data import or mapping errors or savings opportunities. Time spent investigating these changes is generally not wasted. For example, in Figure 20, which baselines year-over-year spend by business unit in a decent-sized bank, it’s important to understand why Consumer and Small Business, Consumer Lending Group, Area Lending, and Electronic Banking increased so significantly.

<table>
<thead>
<tr>
<th>1 to 1 (of 1)</th>
<th>Last Year</th>
<th>Current Year</th>
<th>Y-Y Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retail</td>
<td>$50,792,656</td>
<td>$85,163,590</td>
<td>68%</td>
</tr>
<tr>
<td>CONSUMER AND SMALL BUSINESS</td>
<td>$21,163,542</td>
<td>$19,046,167</td>
<td>-10%</td>
</tr>
<tr>
<td>RETAIL &amp; SMALL BUSINESS BANKING</td>
<td>$13,188,938</td>
<td>$9,219,687</td>
<td>-30%</td>
</tr>
<tr>
<td>MORTGAGE BANKING</td>
<td>$5,889,535</td>
<td>$10,438,729</td>
<td>77%</td>
</tr>
<tr>
<td>CONSUMER LENDING GROUP</td>
<td>$5,569,612</td>
<td>$8,996,184</td>
<td>62%</td>
</tr>
<tr>
<td>AREA LENDING</td>
<td>$1,180,222</td>
<td>$2,650,692</td>
<td>142%</td>
</tr>
<tr>
<td>ELECTRONIC BANKING</td>
<td>$504,701</td>
<td>$514,752</td>
<td>2%</td>
</tr>
<tr>
<td>RETAIL BANK ADMINISTRATION</td>
<td>$287,735</td>
<td>$213,453</td>
<td>-26%</td>
</tr>
<tr>
<td>CUSTOMER STRATEGIES &amp; INSIGHTS</td>
<td>$12,434</td>
<td>$35,927</td>
<td>189%</td>
</tr>
</tbody>
</table>

**Figure 20 A Year-Over-Year Spend Baseline**

**Family the Vendors**

Once all of the fields from the raw data sources have been aligned to a master data cube, the next step is to group vendors into vendor families. This is crucial for meaningful spend visibility, since no analysis or report will be accurate if there are four instances of IBM (referenced as IBM, I.B.M., International Business Machines, and Int. Bus. Mac.). The grouping may also be done on other dimensions as well. For example, it is useful to group together three categories for laptops (laptops, portables, netbooks), or three names for the same division (sales and marketing, marketing, and sales).
The vendor groupings will often be done in a hierarchical manner, as laid out in Figure 21:

![Figure 21 A Vendor Grouping Example](image)

Multiple methodologies can be used in the creation of the vendor grouping. The specific methodology used will be driven by the needs of the business and the analyst. For example, in addition to the multi-level hierarchy illustrated in Figure 21, the organization might choose to segment vendors by negotiation units or legal entities, as illustrated in Figure 22. It is important to understand that there is no "right answer" with respect to grouping vendors – "who owns whom" cannot, for example, be trusted to provide a useful grouping. It is of no value to the negotiation team to group seventeen unrelated companies under the private equity company that owns them. But it very important to group different hotels that are individually owned but are part of a franchise arrangement.

![Figure 22 Segmented Vendor Grouping Example](image)
With respect to grouping vendors, the analyst will have to review each set of vendor records with the same name and different ID. In the example in Figure 23, the vendors should be merged together in the first case as both vendors would be covered by a master FEDEX contract. In the second case, the vendors with the same name are not related in any way and should not be grouped.

<table>
<thead>
<tr>
<th>Vendor Name</th>
<th>Vendor ID</th>
<th>Addr, City, State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fedex Kinkos</td>
<td>435201</td>
<td>142 Middlesex Tpke, Burlington, MA</td>
</tr>
<tr>
<td>Fedex Kinkos</td>
<td>543251</td>
<td>54 Main St; Cambridge, MA</td>
</tr>
<tr>
<td>AAA Locksmith</td>
<td>621342</td>
<td>101 Main St; Medford, MA</td>
</tr>
<tr>
<td>AAA Locksmith</td>
<td>763623</td>
<td>37 Elm St; Albany, NY</td>
</tr>
</tbody>
</table>

**Figure 23 Duplicate Vendor Example**

**Define the Relationships**

The analysis team should define any relationships that are relevant to the analysis and any relationships that support the management and sourcing of vendors. For example, if the organization has a supplier entry for Comfort Suites, Cambria Suites, Mainstay Suites, the Suburban Extended Stay Hotels, and Choice Hotels, but only negotiates master contracts through Choice Hotels, which in this case is actually the parent entity of the other hotels, then the organization would define a family relationship between the identified entities and augment the vendor parenting as appropriate.

In the case of United Technologies, Supply Management might negotiate independently with the different subsidiary companies – Carrier, Otis and UTC Fire – so it may not be wise to put these together into a single family.

An organization may also group instances in the vendor list based on how the organization deals with them. For example, the spend data may include payments to Employees, Government Tax Authorities, and Customers. The hierarchy will need to take this into account. At a high level, it is interesting to know how much Travel and Expense (T&E) spend is paid to employees versus how much T&E spend is paid directly to hotels or airlines[17]. The hierarchical structure allows that next level of analysis. For example, at the top level the analyst can review the mileage reimbursements to employees and then drill down to a lower level to identify the individuals with the highest level of reimbursements.

For government authorities, it’s generally interesting to know the total spending at the top level and then dive down into the different authorities.
How many Vendors Does The Analyst Group?

In the process of constructing a spend cube, some suppliers will get grouped together while others will not. Figure 24 illustrates a fairly typical client situation with approximately 4,000 unique vendor codes in the dataset. In this example, there were 3,835 unique vendor codes. Of these, 333, or 8.6%, were grouped into 188 supplier families and the remaining 3,502 suppliers were not part of any grouping. However, while only 8.6% were grouped, the suppliers that were grouped, which tend to be larger than average suppliers in the dataset, accounted for 35% of the total spend.

![Figure 24 Vendor Grouping Overview](image)

Non-Vendor Hierarchy Changes

Spending data will contain many hierarchies outside of vendors. A normal example of this is the organizational hierarchy – who reports to whom in the company.

Figure 25 is an example of how the analysis team might want to have multiple views of the organization. On the left is the traditional view of the organization where IT groups are spread throughout the different departments. On the right hand side, the different IT organizations are brought together into a new organizational group. Displaying the data both ways will help the purchasing group figure out who they need to work with to influence different parts of the organizational spend.
The ability to reorganize or create different organizational views also becomes important during post merger integration. During the time of the merger, some parts of the realignment of the reporting structures will be clear while other parts will be less clear. To support any post merger integration, the Procurement Group will need to know with whom to work to manage the vendors moving forward.

Some of the most profound insights into spending patterns and savings opportunities are only realized when the analyst is allowed to organize the data across multiple dimensions and organizational schemes. For example, telecom savings are often maximized when the analyst can see which departments in which geographic locations are incurring the most charges, not which supplier is receiving the most payments. Computer equipment savings are often realized only when the analyst can plot the payment data against historical market indices and determine whether or not the best price clause was adhered to. Maintenance costs are often minimized only when the analyst can compare costs for different products across different time periods and determine not only what products have the lowest maintenance costs, but when preventative maintenance can be scheduled to prevent losses associated with downtime. If the organization adopted a single taxonomy such as UNSPNC, it would likely miss out on all of these savings opportunities.

**Define the Commodity Category Structure**

To divide up the spending into the spend categories, every organization is going to need to develop a commodity structure that fits their purchasing patterns. Few companies stick to an industry standard. A recent Aberdeen report\[14\] states that 88% of Best in Class companies use an internally-developed structure. The key thing to understand about a commodity structure is that it needs to be useful to communicate the spending and to organize the groups that will manage the commodities over time.
Some organizations have an existing structure. Some need to develop it. There is no correct abstract structure. The structure needs to work for the organization. Here are two illustrative segments of a commodity structure:

<table>
<thead>
<tr>
<th>Marketing</th>
<th>Advertising</th>
<th>Printing</th>
<th>Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advertising Agency</td>
<td>Magazine, Newspaper</td>
<td>Media Production</td>
<td>Radio, TV, Cable Creative Services</td>
</tr>
<tr>
<td>Direct Mail Forms Lettershop</td>
<td>Lobbying/PR Events, Celebrity</td>
<td>Promotional Mech Trade Shows</td>
<td>Market Research Marketing, Other</td>
</tr>
<tr>
<td>Blanks</td>
<td>Financial Paper</td>
<td>Envelope Printing</td>
<td>Financial Printing</td>
</tr>
<tr>
<td>Paper</td>
<td>Business Forms</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Commercial Print</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual Report</td>
<td>Statementing</td>
</tr>
<tr>
<td>Direct Mail</td>
<td>Warehousing</td>
</tr>
<tr>
<td>Commercial Print</td>
<td>Services</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Facilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
</tr>
<tr>
<td>General Contractors</td>
</tr>
<tr>
<td>Supplies &amp; Materials</td>
</tr>
<tr>
<td>Construction, other</td>
</tr>
<tr>
<td>Architectural Services</td>
</tr>
<tr>
<td>Electrical</td>
</tr>
<tr>
<td>Utilities</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Food Services</th>
<th>Mechanical &amp; Electrical Equip.</th>
<th>Rent</th>
<th>Supplies &amp; Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cafeteria Management</td>
<td>HVAC Equipment</td>
<td>Rent</td>
<td>Plumbing Equip.</td>
</tr>
<tr>
<td>Catering</td>
<td>Mech Elect Equip</td>
<td></td>
<td>Safety Equipment</td>
</tr>
<tr>
<td>Vending, Coffee Svcs</td>
<td>Construction</td>
<td></td>
<td>Signage</td>
</tr>
<tr>
<td>Food Service, other</td>
<td>Management</td>
<td></td>
<td>Warehouse Equip</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Some groups adopt large portions of the UNSPSC or E-Class classification structure, but most either develop their own or make modifications to a general structure. For example, UNSPSC is an OK starting point for the direct goods purchased by a manufacturer and is good for companies sharing transactional data with other companies. The UNSPSC structure, however, is rarely adequate for indirect spend.

As the organization develops its commodity structure, it needs to keep in mind a few guidelines. At the lowest level, a commodity should be large enough to manage, but not too large. For example, the CPO might determine that a commodity category should have at least $250,000 or $500,000 a year in annual spending. Anything smaller will not be managed as a standalone category.
The minimum meaningful amount of spend will determine if the organization breaks out cleaning supplies as a separate category from cleaning. If the cleaning supplies expense is only $25,000 per year, it might not be big enough to warrant a multi-month team effort. Perhaps the team can do a focused negotiation to get a meaningful discount from the vendor.

Another way to think about the right size for a category is to think about dividing the spend base into segments where you can generate a good return on investment from a savings initiative. Let’s assume the cost for a person in the sourcing team is $12,000 a month ($144,000 per year). Your figures may be higher or lower – this is just for illustration. The company will want a good return on this investment. Let’s assume you want a return that is three (3) times the cost of your staff. This means that for each month of effort, the person needs to return $36,000 of benefit. If you have a $2 million category where you expect to generate a 12% savings, you will expect to generate $240,000 of savings. This means your sourcing person can only spend 6 months on this category (6 x $36,000 = $216,000 of “cost”). In this scenario, categories smaller than $2 million will not be fully sourced. This doesn’t include the management time to help this person be successful. Some companies demand more than a 3x return to account for the organizational support for a sourcing effort.

A commodity category in a geography should be large enough to be bid out on a single RFP. In addition, the commodity structure should be designed so a single RFP does not need to span multiple commodities outside of outsourcing or facilities management bids. For example, in the facilities area, an organization might have commodities for cleaning, small repairs, painting, and trash removal, but Supply Management might bid out the entire group in a single RFP. It is useful to have this segmentation since these commodities represent different markets with different vendors.

These examples are just suggestions. There are good reasons for making other choices. Usually, the organization will develop an initial structure and let it evolve over time. As the analysts work with the data, they will determine where the structure is effective and where it needs improvement. For example, should IT Consulting be inside the IT area? Or Professional Services? Or HR? If IT consultants are used as staff augmentation and broadly used across different programming teams, it might make sense to manage them just like regular staff – so put them under HR. If IT consultants are used exclusively on well defined projects with specific milestones, if might make sense to manage them under IT.

There is no absolute “right” answer. Supply Management needs to place a commodity in the hierarchy where it makes sense for the organization, and over time it might change its mind dozens of times. A modern spend analysis system makes this easy – just a few points and clicks, and a few seconds should be all that is required. A sourcing professional’s real concern is that the spending for IT consulting is kept together in a single category, and not spread over different areas.
As the organization begins the mapping process, it may discover new commodities or categories that are currently bundled in with other categories that it wants to analyze separately. A simple example is coffee and water for offices. In some organizations, these expenses are clear and clean and show up under vendors like “Jane’s Coffee Service” or “Poland Springs Water”. Other organizations bundle water and coffee with their office supplies purchases. Still others refuse to pay for these items and require employees to pay for these individually – either using vending machines or bringing in materials from home. So the decision to have a “Water and Coffee” category inside facilities will vary by organization.
06 Define the Starting Rules for Classification and Mapping

Once the data formats are defined, the data loaded, the control totals checked, the indexes defined and loaded, and the vendors grouped, the next step is to determine which transactions are associated with what commodity. This process of classifying the spending into commodities by leveraging knowledge in the existing transaction data is called "mapping."

Use the Secret Sauce

Most companies can do a good job classifying data using the GL Code and the Vendor Name. The vast majority of organizations will achieve an 80% to 95%+ spend mapping by using a “secret sauce” of spend mapping that has been well known by leading spend analysis consultancies for over twenty years. The key is the idea of overlapping spend maps to achieve high accuracy where it matters, while leaving "directional" mappings in place for spend that isn’t addressable anyway, and that you don’t really care about.

Here is how to visualize the process. Imagine a light bulb, and let that light represent your total spending. If you put a piece of Swiss cheese in front of the bulb, the cheese will block some of the light. That piece of cheese represents a mapping rules overlay, and any light that’s blocked represents spend that has been mapped by the rules in that overlay. The "spend" that passes through the holes in the cheese remains unmapped.

If you put a new piece of cheese between the light bulb and the first piece of cheese, the new piece will intercept some of the light before it reaches the first piece. Its rules are therefore "higher priority." Nevertheless, the "spend" that passes through the holes in the new piece is unaffected by its rules. That "spend" may either continue on to hit the first piece of cheese, and be mapped there, or it will pass through both pieces of cheese and remain unmapped.
If you put a third piece of cheese in front of the second piece, it may intercept the spend before it reaches either the second piece or the first piece; and so on.

The overlays "closest to the light" contain more precise rules. Those "farther from the light" contain less precise rules. Precision only exists, and only matters, for some of the spend. The rest is not addressable, and it is uninteresting.

Here is the actual process:

1. Map the GL codes. This is the overlay "farthest from the light," and contains the least precise rules.
2. Map the vendors. This overlay is ahead of GL, so it catches the vendors we care about, and maps them more accurately than the GL mappings.
3. Map the vendor + GL code combinations. This overlay is ahead of the vendor overlay, and it is able to map spending with vendors who provide more than one commodity.
4. Review the mappings and map the exceptions. This overlay is for special rules, such as "We know this GL code is never, ever used except for customer reimbursements," so it is more precise than the previous overlays.
5. Review with knowledgeable users to make necessary corrections

A modern spend analysis system will also have the ability to map phrases in text descriptions. This is necessary when AP data is bereft of information (for example, a hospital buying through a GPO (Group Purchasing Organization) typically records only one vendor, namely the GPO). Text mapping groups should be able to be overlaid with ordinary mappings, as above, to be useful.

Let's go through these steps in detail.
1. **Map the GL Codes.** Many GL codes align nicely with the commodity being purchased. Examples could be “Facilities Cleaning” or “Copier Expense”. Other GL codes do not line up to the commodity. Examples include “3 year fixed assets” or “Other Miscellaneous Other” (not a joke: this is an actual GL code in several organizations). The process involves mapping the GL codes which are meaningful and leaving the rest alone. First, use the spend analysis system to filter the data on a single GL code; then review the vendors, cost centers and descriptions - do they mostly fit into a single category? If the category exists, map it. If the spending is large enough and the category does not exist, create it and map the spending. If the spending is small, map it to the best category that makes sense. Typically, a company will have a few hundred GL codes, so this is a fast process. Note that visibility is required across the dataset while mapping; hence mapping is best performed using the high visibility spend analysis tool itself, not with some offline process.

2. **Map the Vendors.** There are many vendors from which the organization buys a single commodity and some vendors from which the organization buys multiple commodities. For vendors from which the organization buys a single item, the analysts creates vendor rules. For vendors from which the organization buys multiple commodities, it is necessary to determine if Vendor + GL rules will better classify the spending (described below). The general process is to filter on the unmapped spend and start classifying the largest vendors. For most organizations, the top few hundred vendors will represent most of the spending.

3. **Map the vendor + GL code combinations.** For vendors from which the organization purchases multiple commodities, it will be necessary to create rules that segment this spending. For example, spending with IBM with a GL Code for Fixed Assets is likely Hardware, whereas spending in an expense GL Code might be IT consulting or maintenance services. The analysts might make a vendor rule that says Xerox spending is for Copiers, and then create an overlay rule for Xerox that further considers the cost center. If it’s Mailroom, then the commodity is Mailroom services, not Copiers.

4. **Review the mappings and map the exceptions.** Most organizations aim for a 95\% to 99\% classification of spending in the commodity hierarchy. Once the initial mapping is done to the desired level of accuracy, the mapping should be reviewed. Review each Commodity, GL Code and top Vendor. Does each single view make sense? For the Commodity View, do the vendors make sense? For the GL view, do the Commodities make sense? For each Vendor, do the Commodities make sense? It is easy to find errors this way, and equally easy to correct them, if the spend analysis system allows you to do so in real time.
5. **Review with knowledgeable users to make corrections.** Once the internal team is satisfied with the mapping, the mapping should be reviewed with knowledgeable staff in the business units. A simple way to do this is to present a key user with a single page showing them the top vendors for their commodity of interest. Ask them to see if the list makes sense, if there are vendors that shouldn’t be there, or if there are vendors missing.

These reviews serve several purposes. The people working closely with vendors know what they are purchasing from the vendors, so they will be capable of providing the analysts with accurate feedback. Meeting with them will engage them in the process of acquiring useful data. Listening to them and incorporating their input will often begin the process of developing a working relationship.

This process continues until enough of the spend is classified (i.e., not in the unclassified bucket) and each commodity has been reviewed by someone knowledgeable about the company’s use of the commodity. The mapping doesn’t need to be perfect. It will continue to change over time.

It is important that the organization understands that mapping mistakes exist, and that it is trivial to fix them (or, at least, it should be, if you have a good spend analysis system). Otherwise you will face objections from individuals opposed to the spend management initiative (there are always some), buttressed by the "errors" they have found. These objections can be overcome easily, and in front of an audience of executives (so the objections can be dispensed with forever) – by correcting the errors, in seconds, right in front of the objector.

**Don't Forget the 80/20 Rule**

Even today, many spend analysis vendors will make a big deal about classification and the effort involved to cleanse, classify, and group the data to reach the 80% to 90% classification accuracy required to begin analysis. However, the mapping exercise is not very difficult if the platform supports an overlay-type mapping scheme. Even clerical resources with no specific skills or knowledge can successfully map spend using the overlay technique.

Since the top 20% of suppliers and products will account for at least 80% of spend, and since the organization’s buyers will be very familiar with the top suppliers and commodities, it won’t take long to create and vet the first mapping rule overlays. As specific categories are analyzed, rules to map the remaining suppliers and categories can be created. It is always possible to create a new overlay to get more specific in a particular area. Such a mapping strategy will allow even the largest organization to achieve an 80% to 90% classification accuracy within a week – enough to start analysis. Increased accuracy is only required on data subsets that are likely to contain greater savings opportunities.
Much ado is made by some spend analysis vendors of "auto-classifiers" that use custom or third-party databases of Vendors to map spend. Some go further, and use text matching rules, sometimes quite complex ones. These systems have two basic problems:

1) An auto classifier must classify to a fixed classification hierarchy. Typically this is UNSPSC, SIC or NIACS codes, which, unfortunately, are not very good classification systems for indirect spend. Alternatively, the hierarchy may be a custom hierarchy created by a spend analysis vendor. In either case, you will need to build a mapping table from their codes to your structure. There are several issues with these approaches:

- Using a fixed structure means you can’t modify the structure to match to the way the organization actually thinks about its spending. This is quite critical to getting organizational buy-in to the spend map.
- The structure may have sections where it is less granular than you want for your data. You will need to remap the spending after it’s done.
- Vendor name, by itself, is inadequate for a good spend map. What if the vendor provides more than one commodity? And, even if 70% of vendors are classified, who cares? Only a few hundred really matter, and you’re going to have to check those manually anyway, because you can’t afford a mistake. Your first priority for sourcing is not going to be Fred’s Cleaners, so it doesn’t matter if Fred is mapped roughly by GL or more accurately by Vendor.
- Text mapping tends to create long lists of complex matching rules, which are unmaintainable and counterintuitive. It is impossible to figure out which rule is doing what; and after awhile, everyone gives up trying.

2) An auto classifier makes mistakes. They all do; it is inevitable. Every auto classifier’s results must be manually checked for errors, and the manual checking can often take nearly as much time as creating accurate rules, by hand, in the first place. The very first time spending is mapped, there may be a slight time advantage to using an auto-classifier. But when the data are refreshed (next month or next quarter), the auto-classifier will make exactly the same errors that it made the first time, and possibly more. Its output will need to be checked, manually, again. At this point, it becomes more expensive to use an auto classifier than it would have been to create solid rules manually, that never need to change, and that are always accurate.

Many industries have significant “exceptions” to the expected "automated" decision regarding which commodity a vendor supplies, for example:

- Telecom vendors pay churches and schools for cell towers (not for tithes or charity)
- Pharmaceutical providers pay licensing fees to colleges (not for education)
- Payments to banks might be for a T&E card; a lease; a loan payment; or even a legal settlement to a customer.
Most of the effort required for mapping only has to be undertaken once. After the rules are defined, captured, and properly prioritized and layered, they are automatically re-applied every time the master cube or data store is refreshed.

The diagram in Figure 26 illustrates, again, how the different rules interact. As discussed above, the way to think about this is as a series of filters, each of which catches some of the spending and allows the remainder of the spend to be passed through unmapped.

![Diagram of rule overlay](image)

**Figure 26 Mapping Rule Overlay, Visually Explained**

Figure 27 is an example of how three rules can be used to map seven transactions.
There are two rules for Xerox. The first rule (a Vendor rule) maps all the spending for Xerox to Copiers. The second rule (GL plus Vendor rule) has higher priority and maps the segment of Xerox which does mailroom services to the Mailroom category.

The third rule maps all the spending in the GL Rule group for Cleaning to the cleaning category. This moves all the spending for five vendors in the GL category to the Cleaning category in a single rule. In most cases, GL rules map even more vendors. In the future, if the company starts using new vendors and booking this spend to the same GL code, this spend will automatically be mapped to the Cleaning category.

**How much spending gets mapped by different rules?**

The need for multiple levels of mapping is illustrated by analyzing the average percentage of spend correctly mapped when mapping is done by vendors or GL codes alone. A recent review of 12 AP spend cubes from organizations in different industries demonstrated the need for both Vendor and GL mapping to satisfy a 90% threshold for a typical customer. The 90% threshold was used to ignore the “tail” of spending.

To understand if GL codes could be used by themselves to map spending, the team looked at the spending for each GL code. If 90% of the spending for this GL code was mapped to a single commodity, the GL code was declared to be a “single commodity GL code”. If the spending for this individual GL code was split among more than one commodity, the GL code was declare to be a "multi-commodity GL code." The analysts then summed up the “single commodity GL codes” and calculated their spending as a fraction of the total spending. The analysis found that only 30% to 60% of the spending was in “single commodity GL codes” and the rest was multi-commodity codes.
The same analysis was done for the top 500 vendors, and found that only 60-80% of the vendors were mapped to a single commodity. This means that if you only use Vendor rules, you will be “wrong” 20 to 40% of the time. If either vendor or GL code did a good job mapping a cube, the percentages would be close to 100% on one of the axes.

The results are summarized in the diagram in Figure 28.

![Figure 28 Mapping Percentages](image)

For this analysis, only spending mapped to meaningful indirect commodities was considered, e.g. all spending mapped to Exempt or direct categories was ignored. A 90% threshold of spending was used to determine if a GL code or Vendor was eligible for analysis, in order to avoid having vendors with a very small amount of spending (the mostly uninteresting tail) overly influence the outcome.

**Remember, AI is Not Enough**

A number of vendors push Artificial Intelligence (AI) as the answer to rapid cleansing and classification. While the initial sales pitch might sound enticing, one has to remember two things: (1) researchers have been promising AI for over fifty-five years (ever since John McCarthy coined the term in 1956) but have yet to actually deliver; and (2) intelligence is always context-sensitive.
While a significant number of advances have been made in programming, logic systems, database systems, and automated reasoning in the quest to build an AI, AI itself has a long history of failure. Machine translation failed in the sixties. Connectionism was abandoned in the seventies. Japan spent the majority of the eighties trying to build a fifth generation computer using Prolog, which was supposed to be the language of AI; this effort also failed miserably. Expert systems went out of style in the nineties. And the noughts could be the last decade that saw regular research into a revised form of connectionism.

However, even more important, is that intelligence is context sensitive and an AI is generally not capable of determining context. Statistical pattern matching might allow an AI to detect that Int’l Bus Machines maps to IBM, but without guidance, it’s likely that the AI will not be able to accurately determine if “flour. tube” (misspelled “fluor”) is a bakery ingredient or a facilities stock item. Unless a human builds a rule that defines the context, the AI has a good chance of getting it wrong – and, what’s worse, getting it wrong on every single refresh of the data, requiring manual correction each and every time.

A payment to a vendor can mean very different things. Payments to Bank of America could be for a corporate loan (mapped to Corporate Financing); for credit cards (mapped to T&E expense); or for a lease payment (mapped to IT Hardware). Some judgment will be required to determine the right mapping to use.

This is why no vendor can realistically claim that its engine can automatically map an organization’s spend, regardless of how advanced the vendor claims its AI engine is, how many rules are in its rule set, or how many transactions it claims to have mapped. If some of the companies for whom it has mapped data are in the same vertical and have similar spend patterns, there may be a large number of useful rules in the rule set, but a human will still have to review the rule set and select the starting rules. Some use tools to aid in the mapping, but automating the full mapping process is hard. Companies have spent millions on this. IBM has created Watson to play Jeopardy, so amazing things may be possible. If a vendor claims full automation, you should be able to see accurate mapping in an hour or so, for a nominal fee.

If an analyst does not know what a vendor does, the analyst should look them up on the web – “Google” them. If a vendor has material spending and is does not have any web presence, this might raise a bunch of other issues.
Don’t Fall for the Accuracy Myth

One of the myths that continues to be perpetuated is that the mapping requires 100% accuracy to realize the full benefits that spend visibility can provide. In fact, nothing could be further from the truth. A successful spend visibility effort is one that identifies where opportunities lie, and where opportunities do not lie. For the vast majority of spend, categories, suppliers, and business units, a 90% mapping of spend will clearly indicate whether or not there is a savings opportunity to be realized through further analysis and subsequent action.

And sometimes, as per this quote from Jim Polak, PPG’s Director of General Purchasing, in an interview with Global Logistics & Supply Chain Strategies’ Jean Murphy in 2008 [12], less-than-perfect mapping is more than enough:

I have talked to a lot of different companies about this issue and often the mistake people make is in thinking that the data they get from their archives and ERP systems has to be perfect and, unless it is perfect, they can’t do any analysis and can’t hold a sourcing event. Nothing could be further from the truth. It’s true that data enrichment and data cleansing are never ending tasks, but one of the things I stress to people is that they are not doing spend analysis to create pretty charts or to impress people. They are doing spend analysis because they want to see where the money is going and they want to run sourcing events and leverage their spend among fewer suppliers and get savings for their company. And they can do that even with relatively poor data.

If there is no identifiable savings opportunity, then there is no point in mapping the last 20% of spend, because mapping all of the individual exception cases that are contained in the bottom 20% of spend will take 80% of the time. If an organization strives for 100% mapping accuracy on the first pass, by the time the spend is mapped, the cube will be too stale to be of any use to anyone. However, a good analyst can map 80% of global spend in a spend visibility exercise for even the most transaction intensive of the Fortune 500’s in under a week with the right spend analysis tool and get to 95+% in a few weeks.

When the organization reaches a mapping threshold of 90%, it needs to look at the remaining spend that is not mapped. Are any remaining vendors material and interesting? If so, map them. Continue until nothing is obvious to map. Now look at the overall percentage of mapped spend. Has enough spend been mapped? If so, mapping can stop. If not, mapping continues until everything material is mapped or a baseline percentage of spend is mapped. Then the organization dives into analysis. If need be, more mapping rules can always be defined later.

An organization that follows the 80/20 rule will get a much quicker, and a much larger, return on their spend visibility and analysis projects than those that do not. Those who don’t, and try to achieve mapping perfection before beginning analysis, will not only see a lesser return, but risk total project failure, because the risk of failure for any project increases substantially as the implementation cycle lengthens.
07 Avoid the Common Tactical Traps

A spend visibility effort will only yield long-term savings success through strategic spend analysis if it is properly planned and executed and if the design and implementation team doesn’t fall for the tactical traps that are positioned by many vendors and consultants as the quick answer to easy savings.

ERP is Not the Answer

Despite what a few ERP vendors, or their partners, may claim, standardizing on a single ERP system will not resolve the issue of spend visibility. ERP systems were designed for high-level financial analysis and lack the detailed, line-item attribute data needed for any sort of meaningful spend analysis, because they do not contain good commodity information.

The data in an ERP is generally incomplete. It only records part, supplier, and cost, and not invoice, PO, or contract; it does not contain associated transactional data found in P-card systems; and it does not contain any information about supplier diversity, risk, or performance that is required to truly analyze the cost and benefits of historical spending. Furthermore, there will be duplicate vendors, SKUs, and organizational unit identifiers, and, in some cases, even duplicate transactions. Plus, even if the system was utilized to its full potential, and additional information were captured, and if the system was cleansed of duplicates, there is still the fact that ERP data is relatively static and it is generally impossible to modify it. The data will only be as good as the data that was entered by the users. In a system where users were forced to pick a commodity for every purchase, the first item on the commodity list was selected almost 20% of the time, causing the company to think 20% of the spending was for “aircraft supplies” or that 15% of the users were in Afghanistan (since these were alphabetically first on the respective lists). These errors are difficult or impossible to fix, because ERP systems are financial accounting systems, and financial accounting systems resist data modification by design.

The ERP system is focused on the correct financial treatment for every expenditure. If the organization buys one printer, it might expense it. If the organization buys 1,000 printers, it will probably capitalize them. If the organization buys a printer for a customer expense, it will likely book it to a different account code. The purchasing person wants all these expenses brought together. The accounting person requires them to be separated for the correct tax treatment.
Users will also categorize spending by how they think about it, not into classifications based on how commodities are sourced. A user taking a flight for a legal deposition might book the spending to Legal. From their perspective, this is the right way to categorize the spending. "I'm only taking this flight to defend the company in court". This is not what purchasing wants – purchasing wants all the airline travel together in a single category regardless of the reason for the trip.

All that will be accomplished by standardizing on a single ERP and centralizing data to an instance is that the supply team will confirm what they already know – they spend a lot on raw materials with the suppliers they interact with daily. It will not tell the supply team how much they spend on laptops vs. desktops, iOS vs. Android devices, or copying vs. printing; or who is spending the money, and who is using the purchased items.

A Centralized Warehouse Doesn't Solve the Analysis Problem

Some vendors maintain that simply amalgamating all organizational spending data into a centralized data warehouse is enough to meet organizational visibility and analysis needs. In fact, that’s only the first step in a best-case scenario. If this assertion were true, it would mean that the team obtaining the data needs to know – a priori – all the different ways the data might be used and analyzed. Effectively, it’s asking the team to be omniscient. And they are most certainly not.

There are a number of issues with just dumping all organizational data into a centralized warehouse without significant forethought with respect to organization, maintenance, and analytical needs. These include, but are not limited to:

- **Dirty Data** – a significant number of records will contain one or more errors.
- **Missing Data** – many records from the payment systems will only contain the supplier and amount paid and have few or no relevant transaction details; likewise, many invoices from the billing system will be missing corresponding PO or payment details.
- **Duplicate Data** – some records will be duplicated in the P-Card system and AP system; some purchase orders will be duplicated in the ERP and the e-Document management system.
- **No Enrichment** – who are the MWBE suppliers? which suppliers have agreed to adhere to standards of conduct and social responsibility? what are their carbon footprints?
- **Lack of Data Relationships** – how does spend data in the ERP, AP system, and P-Card system relate? how does the inventory and asset data relate to the spend data? how does the purchase data relate to forecast data?
- **Lack of Flexibility** – with data warehouses, change is an onerous and time-consuming process that often takes days or weeks.
Unless these questions are addressed and resolved, the data warehouse will not be of much use. While it may serve as a reference dataset, it will simply increase organizational data sprawl because analysts will make their own copies in Microsoft Access or MySQL and create a blizzard of spreadsheets in an effort to analyze the data on their desktops.

The data warehouse either contains a duplicate copy of the transaction data or a transformed subset of this data. If the warehouse is merely a duplicate copy of the data, then all the warehouse is providing is a place to store redundant data with some different access controls. If it contains a transform of the data, then the analysts will have to hope that the transform was done properly.

When moving data from source systems to a data warehouse, decisions will be made on which fields to keep and which to discard. For example, in Figure 29, System 1 has an accounting code with four subfields – Cost Center, GL Code, Location and Project. System 2 has 6 different fields for the accounting for spending.

<table>
<thead>
<tr>
<th>System 1: Accounting Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>145234 – 1235321 – 0150 - 00</td>
</tr>
<tr>
<td>Cost</td>
</tr>
<tr>
<td>Center</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>System 2: Accounting Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sort Code: A234</td>
</tr>
<tr>
<td>Legal Entity: 101</td>
</tr>
<tr>
<td>Division: 127</td>
</tr>
<tr>
<td>Accounting Group: 57</td>
</tr>
<tr>
<td>Local GL Code: A435 – Gebäude Reinigung</td>
</tr>
<tr>
<td>Global GL Code: 654021 – Building Cleaning</td>
</tr>
</tbody>
</table>

**Figure 29 Accounting Code Examples**
The data warehouse team will need to make a decision on which fields line up with which other fields. They will need to answer a number of questions. For example, is sort code equal to cost center? Is division a location? Does each legal entity have different GL codes? Does each Local GL Code line up uniquely with the Global GL Code? The list of questions goes on. If the analyst takes data from the data warehouse, she has to hope these mappings were all done correctly.

If any data is missing, the spend analysis will need to make use of the original data anyway. In fact, there may be no need for a centralized warehouse at all if the spend visibility and analysis system can extract the relevant data from each system as required, normalize it to a common format, fix and enrich the dirty and missing data via user-defined rules, and create the required relationships. In fact, in most cases spend visibility requires visibility into all relevant organizational data, not just a duplicated subset of the data in some centralized data warehouse.

A Common Taxonomy Only Creates New Problems

While some vendors push the centralized data warehouse, because that’s the solution they have, others push a common taxonomy, such as UNSPNC. A common taxonomy can provide a starting point for relating all organizational data relating to products and services, but it is not sufficient for true spend analysis.

The UNSPSC code structure is trying to be all things to all people. As such, it’s great for some uses and poor for others. It’s easy to poke holes in the structure. Simply go to the UNSPSC web site, http://www.unspsc.org/, and download the current structure. For example, the authors downloaded V13.1201 This is a 3,511 page listing of the current structure at the 8 digit code level. In this version, 151 pages (3% of the 3,511 pages) related to live, fresh cut or dried roses. This is wonderful if the organization is in the rose business. But this is going to be viewed as less than useful by the average user. If nothing else, the organization will need to trim out all the segments that are not meaningful to its users. So, since the organization has to modifying the structure before it even begins its first analysis, why not modify it so that it actually meets the users’ needs?

Unless the goal is to inspire a UFC Grudge Match in the hallway between the analysts to decide who gets to implement his or her changes next, any attempts to settle on a single taxonomy (for a single “master” cube) will be a waste of time. The reality is that analysis cubes and reports need to be private, because analysts will often need to slice, dice, and taxonomize their copy of the data eight ways from Sunday until they find the golden needle of insight in the yellow transactional haystack. Often, only the final cube or report that illustrates the savings opportunity should be shared, and only in a read-only mode.
Figure 30 presents an example of the UNSPSC structure, a taxonomy which contains over 50,000 commodities. For some situations, this can be extremely useful. In other situations, the large number of different codes confuses matters by dividing spend that the analyst believes should be in the same bucket into a large number of discrete buckets. For example, a purchase of a pen and a pad of paper from an office supply store could end up in very different parts of the hierarchy. This goes against the general idea that similar items should end up near each other in the commodity structure. "Nearness" is a critical outcome for a successful sourcing analysis.

In the UNSPSC structure, paper could appear under many different codes, scattered through the UNSPSC hierarchy:

14000000 - Paper Materials and Products
44000000 - Office Equipment and Accessories and Supplies
55000000 - Published Products
60000000 – [...] Educational Equipment and Materials and Accessories and Supplies
82000000 - Editorial and Design and Graphic and Fine Art Services

<table>
<thead>
<tr>
<th>Commodity 127,012,284 of 127,012,284</th>
<th>▼ Amount</th>
<th>▼ Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼ General &amp; Administrative</td>
<td>86,840,687</td>
<td>169,767</td>
</tr>
<tr>
<td>▼ Office Supplies</td>
<td>44,900,569</td>
<td>113,669</td>
</tr>
<tr>
<td>▼ General Office Supplies</td>
<td>44,490,573</td>
<td>113,360</td>
</tr>
<tr>
<td>▼ Copier Paper</td>
<td>409,965</td>
<td>279</td>
</tr>
<tr>
<td>▼ Office Equipment</td>
<td>30,524,722</td>
<td>42,267</td>
</tr>
<tr>
<td>▼ Distribution</td>
<td>7,749,318</td>
<td>3,866</td>
</tr>
<tr>
<td>▼ Office Administration</td>
<td>3,566,089</td>
<td>9,945</td>
</tr>
<tr>
<td>▼ Marketing Advertising</td>
<td>40,171,557</td>
<td>29,568</td>
</tr>
<tr>
<td>▼ Printing &amp; Direct Mail</td>
<td>40,171,557</td>
<td>29,568</td>
</tr>
<tr>
<td>▼ Financial Printing</td>
<td>27,927,236</td>
<td>6,470</td>
</tr>
<tr>
<td>▼ Business Forms</td>
<td>9,800,916</td>
<td>6,857</td>
</tr>
<tr>
<td>▼ Commercial Print</td>
<td>1,732,240</td>
<td>16,024</td>
</tr>
<tr>
<td>▼ Direct Mail Forms</td>
<td>711,204</td>
<td>237</td>
</tr>
</tbody>
</table>

Figure 30 UNSPSC Example

In a more natural coding structure, the paper spending is grouped under either Office Supplies or in the Printing and Direct Mail category.

Some organizations group their spending by business unit. These organizational taxonomies are not always appropriate for certain types of spend analysis. Consider the situation where an analyst wants to analyze direct mail spend, where direct mail consists of lists of names that are printed on letter head and stuffed into envelopes with marketing materials. If the spend is classified by business unit, where is it? In some organizational structures, the spend ends up in four (or more) different business units and gets lost. Consider the example in Figure 31.
To find all the spending, the analyst will need to obtain the spending from all these different business units. If the analyst wants to benchmark the rates proposed by a vendor, these different business unit managers will need to be queried. If direct mail is a major category, the commodity organization needs to support this and make this analysis easy.

**Crosstabs Aren't the Answer Either**

A crosstab is an N-dimensional table that plots spending on one dimension against one or more other dimensions, such as department spend against technology spend by category. While this useful, it’s not sufficient for complete analysis. For example, while it may be interesting that IT spends the most on hardware and Human Resources spends the least, that is to be expected. IT is very hardware intensive, HR is very human intensive. Secondly, it says nothing about whether or not either bucket of hardware spend has any savings opportunities. What is the average cost per machine? How does that cost compare to market average? Is there a common supplier that could be used for volume leverage?

Crosstabs, and the pivot tables that underlie them, are most useful when built in large clusters. There should be dozens, if not hundreds, of them booked automatically down some dimension of interest (like cost center by vendor, booked by commodity). They should also be created dynamically, ready for insert into existing models, with the raw data available for secondary and tertiary analysis to explore trends and calculate real savings opportunities.
Furthermore, pivot tables alone are not very useful in the presentation of multidimensional data. That’s why the original Mitchell Madison Group’s (MMG) commodity spend report book top GL accounts, top Cost Centers, and top Vendors by Commodity, with a monthly split of spend showing the share of total category spend at each display line.

A clustered crosstab, like the one developed by the MMG, is much more useful because it supports multiple views and insights and books on multiple dimensions, including GL accounts, Cost Centers, and Vendors by Commodity. For example, the MMG report in Figure 32 allows an analyst to more quickly get a feel for high-level spending patterns in the organization.

Furthermore, the multi-page MMG report, built ideally with a simple data set extract, is best inserted directly into a user-defined Excel model that allows the user to alter the output to meet her needs, putting true analysis directly in her hands. For example, the report could be just as easily booked by Vendor, showing Commodity, Cost Center, and so on. It’s this ability to support multiple views and insights that allows the user to quickly identify real savings opportunities instead of just large spend buckets.
To understand why a crosstab alone is not enough, consider the example in Figure 33 where an analyst is trying to understand the Retail department’s use of Keane. If the analyst started with a crosstab of 2010 spending by vendor and business unit, she would quickly see that of the 13.8 million of total spend, Keane represented 2.3 million or 16% of the total spend and that Retail accounted for 0.7 million or 31% of the total spending on Keane.

![Figure 33 Vendor Spending By Organizational Unit Crosstab](image-url)
However, the analyst is left with as many questions as there are answers. For starters, is the spending consistent over time or does it represent a defined project? To answer this question, the analyst would then have to do a crosstab of quarterly spending for the past few years to see that the spending is not steady over time, that most of the spending is in the last quarter, and that the spend is growing, as in Figure 34.

![Figure 34 Quarterly Vendor Spending by Organizational Unit Crosstab](image)

The analyst might also surmise that the annual spending might now be well over 2 million and that the spending is spread over different areas and rates. To answer these questions, knowing that Keane is a provider of contract labor to the organization, the user would then build a crosstab by facility and rate range and find out that all of the facilities are now using the service provider and most are using resources in the $55 to $60 an hour range, as per the example in Figure 35.

![Figure 35 Facility by Contractor Rate Range Crosstab](image)
Only now would the analyst realize that a significant savings opportunity lies in getting reduced rates for the Retail resource classes that fall in the $55 to $60 range. Without this deep dive, the user probably would have assumed that the largest savings opportunities lie in reducing the payments to Adecco from Operations and Headquarters and to Keane from Operations, and overlook this savings opportunity.

08 Define the Cubes

Once the schema has been defined, the vendors have been familied, and the transactions have been mapped to a useful structure, it’s time to define the cubes that will be used for visibility, analysis, and reporting.

Define the Base Cubes

Note that the goal is not to define every cube or report that might ever be used, but simply those cubes that will give the organization regular insight into spend by category and commodity, vendor, department and geographic unit as a starting point. A leader in spend visibility and analysis will build hundreds of cubes and reports over the lifetime of a solution, and throw the vast majority of them away, keeping only those that have proved to be truly insightful, and only for as long as they are relevant for analysis and spend tracking.

The diagram in Figure 36 illustrates some basic spend cubes that an organization might build and the uses they serve. The number of cubes built by an organization in the analysis of its spend should only be restricted by the imagination of the analyst. Some leading Procurement organizations maintain over 75 cubes in the pursuit of greater cost savings and efficiencies.
Just like in the main AP cube, when the analyst creates different cubes for different areas, she needs to ensure that the spending in the different cubes is complete or know where it is different. If the detailed cube represents the spending for one vendor, the totals are easy to check against the overall spending with that vendor.

The importance of data validation cannot be understated. For example, one of the authors asked a major PC supplier for transaction level detail from client invoices for the last year. The PC supplier provided a report the next day, but when total spending was compared to known total spending with the vendor, it was discovered that the detail file was much smaller. It turns out that several business units, operating under a different trade name, were not included in the report. The names of the missing units were provided to the sales rep who was able to provide the missing data in a matter of hours. But more interestingly, when the data was provided, it was found that the extra divisions were not getting the corporate discount on their spending. This resulted in an immediate credit back from the vendor.

**Build Enriched Cubes**

After the base cubes have been constructed, the data can be enriched with many different types of data that may be collected in the organization. The ability to coordinate different types of data can add significant value to the organization.

**Channel Compliance**

Figure 37 is an example of an internal compliance report by business unit that illustrates how compliant each business unit is with corporate procurement directives (designed to insure regulatory compliance) that dictate the proper channels to be used for the acquisition of goods and services.

In this example, a determination was made for each commodity if POs were required for the commodity. The spending in these commodities is summed by Business Area to determine the possible PO spend. The actual spending done via PO is then compare to this figure to determine the channel compliance rate.
Many companies have systems for electronic bidding, contract management, purchase orders and payments. Companies should understand how much of their spend is managed. Spend analysis allows users to see their spending across all of these different systems.

In a theoretically perfect world, one would identify that a particular payment on an invoice was made to a vendor and that it matched a valid PO in the purchase order system, an active contract in the contract management system, and an RFP in the e-RFX system that led to the contract award. However, at most companies, which will have four different systems, the systems are not linked. So this ideal scenario is the exception rather than the rule, because in the normal case transactions cannot be linked between the systems.

However, with a good spend analysis tool, it is possible to trace spending with a single vendor across the different systems, albeit with some effort and ingenuity. The issue is typically that there will be different vendor IDs if there are multiple AP systems. Furthermore, even in the case where there is a single AP system, even though the PO system could be connected and use the same vendor master, this typically is not the case at many organizations. And, most contract management and e-RFX systems maintain their own vendor masters, since a vendor can be created by anyone on a moment’s notice. These systems have no restriction on the creation of a vendor because this needs to be done quickly. For example, a user might need to file an NDA with a vendor that the organization has never done business with, and in the e-RFX system to allow the vendor to bid. Any attempt to use the existing master will typically be blocked by Accounts Payable, since they will want to limit their vendor master to vendors the organization has approved for payments, in an attempt to limit fraud and/or simplify record keeping. As a result, the vendor masters for these different systems will never be in sync.
Since the vendor masters for the different sourcing and procurement systems will never be in sync at most organizations, the spend analysis system needs to be able to collect data from these different systems and bring it together for analysis. And while this will often require some elbow grease on the first attempt to integrate the transaction files, the effort will usually pay off in spades. Consider the example in Figure 38.

**Figure 38 A Combined Spend View**
In this example, there is a large amount of disparity across the multiple systems. The contract in place for $300,000 must have bypassed the e-RFX system as the total from the e-RFX system is $75,000. If organizational policy is to bid out all contracts over $100,000, then someone is bypassing organizational policies and likely costing the organization real dollars.

**Diversity**

Many organizations have goals to do more business with vendors owned by historically disadvantaged groups. In the US, these include business owned by minorities, women, and disabled veterans. The organization needs to augment its spending data with minority and women-owned business enterprises (MWBE) information in order to not only report on its diversity information, but verify that it is meeting any obligations it has with respect to governmental funding or that it meets the requirements necessary to apply for MWBE benefit programs currently being offered by a government entity.

Some organizations will collect this data directly from their vendors when they set up the business relationship with the vendor. Some organizations will engage a third party (such as Dun and Bradstreet; CVM Solutions; and others) to match their vendor list against registered diverse suppliers. This matching should be done in a smart manner to minimize the costs. This data will also need to be maintained over time.

Figure 39 is an example of a MWBE report that reports on the number of diverse suppliers, diverse spending vs. spending targets, and MWBE spending goals.

![Diversity Reporting](image)

**Figure 39 Sample MWBE Report**

The report in Figure 40 breaks down the MWBE spending goals by operating committee member.
Risk

An organization might track financial information associated with key vendors as well as other information that can identify current and future business risks so that informed spending decisions can be made, taking into account not only current costs but potential losses associated with doing business with a riskier supplier over the long term\[15\]. Otherwise, an analyst might make a decision that saves the organization a million dollars today but costs the organization ten times as much in the future when a sole supplier runs into difficulty.

The business impact report in Figure 41 illustrates the current percentage of the supply base that is high risk and the percentage of total organizational spend that is associated with those suppliers. In this example, 17% of organizational spend is with 17 high-risk vendors.
This information takes significant effort to collect, so not all organizations do this. In this example, a manufacturing company keeps track of vendors supplying materials or components that are incorporated into their finished products. This data is collected for each relevant vendor by an in-house system and is linked into the spend cube by an index on the vendor ID.

**Compliance**

If the organization tracks information on suppliers, this data can easily be linked to the spending data\(^{[18]}\). Data collected might include financial solvency, on-time delivery, local-market stability, or compliance with social or environmental initiatives. Once the data is linked with the spend data, then the organization will be able to easily compile supplier rating by business unit, providing insight as to which business units are performing better than others and which need more attention. This is not necessarily a substitute for a full-fledged Supplier Information Management or Supplier Performance Management system; but it is useful to see SIM and SPM information in the context of the spend cube.
09 Analyze, Assess, Report, Decide

Once the data are enriched, a full-fledged spend analysis effort can begin. This effort, only possible because a well thought-out effort was undertaken to enable visibility into the full breadth of the organizational spend, starts with a tactical analysis to identify immediate payback opportunities. It then moves on to analytic efforts designed to identify low-hanging fruit savings opportunities. Once the immediate savings opportunities are exhausted and the low-hanging fruit has been picked, the focus shifts to strategic, long-term savings opportunities identified through a sophisticated analysis that goes beyond simple AP spend to make use of the non-ERP data and the data enrichments added in the last phase.

Getting the organization to take action is critical to the success of the spend analysis effort. If the organization doesn’t do anything with the data, the effort was for naught.

Find Opportunities in the Data

Once the core AP Cube has been created, companies typically review the data for new savings opportunities, such as those summarized in Figure 42. This represents a commodity by commodity review of the spending, ranking each commodity on the size of the opportunity and degree of difficulty to achieve these savings.

Review Process for Each Category:

- **Current status.**
  - Is there a program in place for this commodity?
  - Are there contracts in place with key vendors?

- **Review spending.**
  - Trend line (growing, flat, shrinking)
  - Shift in product mix expected (either demand or supply)
  - Improved mappings

- **Review opportunity.**
  - Estimate type and size of opportunity for sourcing, benchmarking, and other savings initiatives
  - Estimate difficulty of pursuing initiative.

- **Other factors for the commodity.**

Here are some different types of savings programs that an organization might pursue:
<table>
<thead>
<tr>
<th>Opportunity</th>
<th>Definition</th>
<th>Source of Savings</th>
</tr>
</thead>
</table>
| Sourcing         | Savings program involving a go-to-market in a competitive bidding situation. May involve RFXs, auctions, and/or decision optimization.                                                                        | + Market competitiveness  
|                  |                                                                                                                                                                                                              | + New suppliers  
|                  |                                                                                                                                                                                                              |   - Cost of changing vendors (typically 2%-4%)  
|                  |                                                                                                                                                                                                              |   - Disruption in organization  
|                  |                                                                                                                                                                                                              | + New technologies  
| Invoice Review   | Complete review of contracts and all invoices to ensure contractual prices have been charged in the past and that future pricing is no greater than benchmark. Review items bought to see if all items (or service levels) purchased still needed. | + Savings on past spending from errors and omissions  
| Benchmarking     |                                                                                                                                                                                                              | + Future savings from reduction of pricing to benchmark  
|                  |                                                                                                                                                                                                              | + Fast implementation – no change in vendor or program  
| Consolidation    | Category where commodity spending needs to be consolidated to one or more vendors. May be an opportunity for benchmarking.                                                                                    | + Reduction of pricing to level of new contract  
| Demand           | Savings program focusing on reducing demand for a good / service                                                                                                                                              | + 100% savings for each unit no longer required  
|                  |                                                                                                                                                                                                              | - Not easy in all categories  

**Figure 42 Different Approached for Generating Savings**

A tangible output of the review is to have the commodity managers document the opportunities in each individual category. This can serve as a template for documenting opportunities in other categories. A simple one page overview captures the key information on the spending in the category; the opportunities; any barriers; the plan to achieve the savings; and who needs to be involved to capture the savings, both inside and outside of procurement.

Figure 43 is a template for an individual category opportunity assessment:
The data on each category is then collected, aggregated and summarized. Companies will then typically begin to address the first wave (Figure 44) of projects drawn from the categories which have a high return and a (relatively) easy implementation.
While a modern spend visibility system is an order of magnitude less expensive than it was even five years ago, like any software-centric solution, there are direct and indirect costs associated with the solution. As a result, there is a need to realize an ROI as soon as possible to justify the expenditure and maintain support for the ongoing analysis.

The best way to realize a quick payback is to start with those projects that are most likely to result in immediate cost reductions to the organization and, if feasible, refunds for large volume purchases or overpayments. The best way to find these funds is through one of the following four analysis efforts.
Best-Price Contracts

A quick win can be found in “best-price” contracts, especially if the organization buys a lot of computer and electronics equipment. Most providers will happily agree to a “best-price” cost, but many bill the organization a rather flat rate for the same equipment over time. Considering that computers and associated electronic categories tend to depreciate about 2% to 3% a month, there’s no way that the price should be flat over the course of a year – which is precisely what most spend analysts find when they plot pricing. A quick comparison of amounts paid with a historical market index will often reveal about some overpayments due back to the organization. A recent Aberdeen study\textsuperscript{14} indicates this ranges between 0.1 and 0.5% of the payments.

Figure 45 presents a disguised best-price analysis on computers and computer components that a large organization acquired at “best price”. This organization was overcharged somewhere between 12.6%, which is the straight line decrease between price paid and “best price” indices, and 25.7%, which is the external “best price” benchmark; either one represents a significant savings on the 1.6 million of spend being analyzed.
Substitutions

In certain commodity verticals, like technology, the vendor will often lowball a product that it knows to be at the end of its production life. Then, three to six months later, when the product is unavailable, it will substitute a significantly higher priced product, knowing that the buying organization is unlikely to catch the unapproved price increases because of a change of SKU.

However, if the organizational buyers had the foresight to include a standard substitution clause – one that mandates that any substitutions must be for a functionally equivalent item at a cost equal or less than the agreed upon cost – then this should not happen. If this substitution clause is resisted, then the contract should, at the very least, require that you agree to the changes.
For many mid-sized or large organizations that buy a lot of traditional office supplies, this often represents a cost recovery opportunity of 10% to 20%. Some specialty cost recovery consultancies make their living analyzing spend in the electronics and office supplies verticals, because they will typically find millions of dollars in overpayments and discounts that have not been honored.

The best way to find these unapproved price increases is to create a Top-N report of off-contract item purchases by spend, and the best way to estimate the magnitude of the opportunity created by off-contract items is to plot the % of off-contract purchases over time. When the rate of non-compliance exceeds a certain threshold, the savings opportunity will usually be significant.

Figure 46 illustrates a plot of the number of off-contract purchases over time and the magnitude of the potential savings opportunity from eliminating substitutions and getting more items under contract.
The report in Figure 47 identifies the Top-N off-contract item purchases by total spend.

### Figure 47 An Off-Contract Purchase Assessment by SKU

#### Service Contracts

If the organization uses a lot of temporary labor or consultants, these contracts often represent a prime opportunity for a quick cost reduction. Most organizations typically pay 10% to 30% more than they need to with respect to the function that is being filled, as a detailed spend analysis against market pricing will quickly reveal. In fact, in some areas, the savings potential is so significant that there a number of consulting and solution providers that specialize solely in temporary labor management.

For geographically-distributed organizations, sometimes the best strategy is to use a mix of national and local placement agencies and consulting firms to ensure that the organization is always able to get the people with the right skills at the best rates. While a nationwide staffing vendor might be able to amortize overhead costs and provide the organization with resources at a lower hourly rate, if these resources are not local, after (travel) expenses are included, the all-in rate can be significantly higher than just using local resources. For example, one solution provider released a case study showing that their client was spending 25% more to fill their temporary labor and consulting positions than they needed, which translated into a savings of $20M[4].
One of the best ways to reduce labor costs is to identify the rate bands that are costing the organization the most. The report in Figure 48 breaks down temporary labor costs for a single class of temps by facility and rate band and allows the organization to quickly determine that its greatest savings opportunities lie in reducing the cost of resources in the 25-30, 30-35, and 20-25 rate bands.

<table>
<thead>
<tr>
<th>Facility</th>
<th>15 to 20</th>
<th>20 to 25</th>
<th>25 to 30</th>
<th>30 to 35</th>
<th>35 to 40</th>
<th>40 to 45</th>
<th>45 to 50</th>
<th>Grand Total</th>
<th>Average Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oak Street</td>
<td>38,731</td>
<td>894,332</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>933,063</td>
<td>19.73</td>
</tr>
<tr>
<td>State Place</td>
<td>332,683</td>
<td>799,342</td>
<td>579,908</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,711,934</td>
<td>20.06</td>
</tr>
<tr>
<td>Green Center</td>
<td>250,599</td>
<td>224,383</td>
<td>374,172</td>
<td>51,165</td>
<td></td>
<td></td>
<td></td>
<td>900,325</td>
<td>20.19</td>
</tr>
<tr>
<td>Village Center</td>
<td>76,559</td>
<td>652,148</td>
<td>317,981</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,046,688</td>
<td>20.76</td>
</tr>
<tr>
<td>Maple Street</td>
<td>181,185</td>
<td>47,363</td>
<td>433,087</td>
<td>35,610</td>
<td></td>
<td></td>
<td></td>
<td>697,244</td>
<td>21.16</td>
</tr>
<tr>
<td>Maroon Center</td>
<td>153,812</td>
<td>282,270</td>
<td>667,365</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,103,447</td>
<td>21.61</td>
</tr>
<tr>
<td>Redwood Street</td>
<td>36,498</td>
<td>524,152</td>
<td>467,884</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,028,533</td>
<td>21.72</td>
</tr>
<tr>
<td>State Center</td>
<td>57,702</td>
<td>45,445</td>
<td>845,086</td>
<td>173,341</td>
<td></td>
<td></td>
<td></td>
<td>1,121,573</td>
<td>24.54</td>
</tr>
<tr>
<td>Birch Street</td>
<td>57,048</td>
<td>557,519</td>
<td>1,326,627</td>
<td>819,555</td>
<td>25,080</td>
<td></td>
<td></td>
<td>2,785,829</td>
<td>24.70</td>
</tr>
<tr>
<td>Purple Center</td>
<td>25</td>
<td>360,114</td>
<td>2,070,829</td>
<td>718,461</td>
<td></td>
<td></td>
<td></td>
<td>3,149,429</td>
<td>25.24</td>
</tr>
<tr>
<td>City Center</td>
<td>3,180</td>
<td>476,655</td>
<td>65,726</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>545,501</td>
<td>25.47</td>
</tr>
<tr>
<td>Park Hill</td>
<td>106,601</td>
<td>353,678</td>
<td>267,579</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>727,859</td>
<td>25.63</td>
</tr>
<tr>
<td>Black Center</td>
<td>4,032</td>
<td>275,143</td>
<td>91,018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>370,192</td>
<td>25.99</td>
</tr>
<tr>
<td>Yellow Center</td>
<td>289,905</td>
<td>142,603</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>432,510</td>
<td>26.45</td>
</tr>
<tr>
<td>Town Center</td>
<td>153,992</td>
<td>587,245</td>
<td>1,018,173</td>
<td>59,923</td>
<td></td>
<td></td>
<td></td>
<td>1,819,332</td>
<td>27.22</td>
</tr>
<tr>
<td>City Plaza</td>
<td>20,373</td>
<td>119,163</td>
<td>350,800</td>
<td>92,240</td>
<td></td>
<td></td>
<td></td>
<td>582,576</td>
<td>28.56</td>
</tr>
<tr>
<td>State College</td>
<td>65,877</td>
<td>883,427</td>
<td>308,555</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,257,900</td>
<td>30.76</td>
</tr>
<tr>
<td>Grey Center</td>
<td>338,636</td>
<td>164,516</td>
<td>107,400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>610,552</td>
<td>32.70</td>
</tr>
<tr>
<td>Brown Center</td>
<td>94,121</td>
<td>135,982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>230,103</td>
<td>32.77</td>
</tr>
<tr>
<td>American Plaza</td>
<td>85,682</td>
<td>125,541</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>211,223</td>
<td>32.78</td>
</tr>
<tr>
<td>Village Place</td>
<td>17,854</td>
<td>95,494</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>113,348</td>
<td>34.10</td>
</tr>
<tr>
<td>White Center</td>
<td>60,653</td>
<td>135,743</td>
<td>400,889</td>
<td>109,606</td>
<td></td>
<td></td>
<td></td>
<td>706,890</td>
<td>37.68</td>
</tr>
<tr>
<td>Grand Total</td>
<td>1,252,791</td>
<td>4,990,795</td>
<td>9,962,516</td>
<td>5,162,217</td>
<td>1,143,114</td>
<td>508,280</td>
<td>109,606</td>
<td>23,129,329</td>
<td>24.31</td>
</tr>
</tbody>
</table>

Figure 48 Labor Cost by Facility and Rate Band

The savings opportunity for the identified rate bands becomes even more apparent if the data is plotted graphically, as per the diagram in Figure 49 where the indicated rate bands clearly take up the most area.
Once the critical rates have been reduced, the best way to further reduce temporary/contract labor costs is with the right mix of local, national, and international resources. The analysis in Figure 50 illustrates the effect of mixing onshore and offshore at one organization that decided to outsource some of its graphics work. With the proper mix of onshore and offshore resources, the organization was able to reduce its cost by 25%.
Figure 50 Effect of Mixing Onshore and Offshore Labor

Note that it is easy to do a spend analysis against contract pricing without “integrating” the spend visibility solution with a contract management solution. The two-step process of integrating contracts information into the spend analysis system is straightforward:

- Create a "contract" dimension in the spend analysis system that contains a unique identifier, such as name or ID, for each contract.
- Define rules to map the potential spending to each contract making use of the fact that spending on a contract is a function of supplier, commodity, and date range (e.g. Kelly Services, Network Analyst, January 2010 – June 2010 would map to the Kelly Services Labor Contract for 2010).

Note that the resulting "contract" dimension merely identifies spending that "ought to have been" or "might have been" on contract. However, any spending that isn’t associated with a contract must be maverick (off contract) spend.
**Duplicate Payments**

Duplicate detection goes beyond vendors and includes orders, invoices, and receipts. The detection of duplicate invoices can be very important, because duplicates are a major cause of overpayments in many organizations. In the examples in Figure 51, the first example is an exact duplicate. In the second example, the records appear to be duplicates with a different entry date. The third example is likely a recurring invoice with the same amount each month, but should be investigated to verify (since some organizations will resubmit an invoice after 30 days if not paid).

<table>
<thead>
<tr>
<th>VENDOR</th>
<th>INVOICE #</th>
<th>Date</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIBERTY COMPUTERS</td>
<td>L004501</td>
<td>12-Apr-11</td>
<td>$15,075.00</td>
</tr>
<tr>
<td>LIBERTY COMPUTERS</td>
<td>L004501</td>
<td>12-Apr-11</td>
<td>$15,075.00</td>
</tr>
<tr>
<td>PEABODY PRINTING</td>
<td>10032</td>
<td>12-Jan-11</td>
<td>$5,645.00</td>
</tr>
<tr>
<td>PEABODY PRINTING</td>
<td>10032</td>
<td>5-Mar-11</td>
<td>$5,645.00</td>
</tr>
<tr>
<td>GLORIA CLEANING</td>
<td>789</td>
<td>10-Jan-11</td>
<td>$2,172.00</td>
</tr>
<tr>
<td>GLORIA CLEANING</td>
<td>912</td>
<td>10-Feb-11</td>
<td>$2,172.00</td>
</tr>
</tbody>
</table>

**Figure 51 Duplicate Payment Example**

Any (suspected) duplicate should be researched thoroughly. Systematic problems (such as an invoice being loaded twice) need to be corrected promptly. Regular data entry errors in processing indicate the need for process improvements. Intentional fraud, always a possibility, must also be promptly identified and stopped.

**Move on to Low-Hanging Fruit Opportunities**

Once all of the immediate payback opportunities have been identified, the next step is to use spend visibility to identify low-hanging fruit opportunities that can lead to immediate cost reductions going forward. The following three opportunities are usually the most useful at the average organization.
**Off-Contract Spend**

When one looks at the summary information for the vendors, one can see the total amount of spending and the amount done via POs. If a goal is to move more spending to PO's, Supply Management would focus on vendors where there is partial PO usage. In the illustration in Figure 52, one can see the calculation of “Percent PO” which is the portion of spending which is PO based.

![Figure 52 Percentage of Vendor Spend on PO Summary Report](image)

This analysis could also be done by organizational unit. The analysis in Figure 53 shows how two organizations are making significant improvements to their use of PO's while other units are not making significant changes.
Chances are that the organization has a lot of spend for which there are no contracts in place which specify the overall unit pricing or pricing terms. While it’s likely the case that some of this spend should be off-contract because it’s typically cheaper in the long run to spot-buy in some commodity markets than to enter into long-term contracts, there will likely be significant buckets of spend where the organization could get better prices by amalgamating the spend and leveraging the volume using an organizational contract. Since there are no contracts currently in place, the organization can effect the contracts immediately and start realizing savings right away.

The graphic in Figure 54 illustrates the potential savings at a 10% discount for an organization with 326 million in spend that has 75 vendors without contracts that receive over $300,000 in annual spend individually and over $108 million, or 33% of total spend, in aggregate.
Maverick Spend

After the organization has leveraged its previously unleveraged volume to achieve immediate price reductions, the next step is to tackle the issue of unrealized savings. In an average organization, up to 40% of negotiated savings from a strategic sourcing event will never be realized. There are a number of reasons for this, including unanticipated demand patterns that result in increased logistics costs due to expediting, higher than expected inventory management costs, unexpected increases in logistics costs (due to availability and fuel surcharges), or interruptions in supply due to natural disasters, work stoppages, and raw material or labor shortages. However, the biggest loss is due to non-compliance. In fact, an Aberdeen study found that in an average organization, only 64% of purchases were on contract\textsuperscript{[5]}. That says that 36% of purchases were off-contract or not compliant with contract terms and conditions.

An average organization can lose millions of dollars to maverick spend, even if the purchases were below contracted rates. Instead of low per-unit costs, many manufacturers will instead offer volume-based discounts, which are only credited to the buying organization when a sufficient volume of goods has been purchased. Thus, if the buying organization commits to buying eight hundred thousand units, but only buys six hundred thousand units and the discount threshold was seven hundred thousand units, the buying organization will never see the 10% discount, which could translate into over a million dollars in savings if each unit was fifteen dollars. Thus, the buyer who thought he was doing the organization a favor by buying the product from a vendor down the street who offered a 3% discount on one hundred thousand units, effectively cost the organization over nine hundred thousand in savings.
Thus, it’s very important for the spend analysts to determine if there is rampant maverick spend in any category where it could have a significant impact on negotiated savings, trace the source, and put a stop to it immediately. Then, the analyst can put in place reports that identify and trace the remaining maverick spend and send those reports to the supervisors of the worst offenders every month to curb the practice over time. (While it is true that all maverick spend has a cost, in line with the 80/20 rule, it’s not worth the analyst’s time to try and stop all the drips after the geyser has been plugged, because there are bigger cost savings opportunities to be attacked.)

It’s easy to determine whether or not there is enough maverick spend in a category to make a spend reduction effort in that category worthwhile. Simply:

- Create a report in a category of last year’s spending by vendor.
- For each vendor, indicate if there exists a contract with specific rates for the goods or services being purchased.
- Calculate what fraction of the spending does not have contracts in place. This is the Bypass Spend Amount.
- Multiply the Bypass Spend Amount by the estimate for how much is saved using the contracted vendors. This is the total opportunity.

If the bypass spending is greater than some limit (perhaps $25K), then the category should go into the short-term opportunity queue, as it will only take an analyst at most a day with a good spend visibility system to track down the worst offenders, create a report highlighting the opportunity, and distribute it to the appropriate managers to take action.

The report in Figure 55 calculates bypass spend by vendor and the potential savings that would be achieved by shifting the bypass spend to approved, contracted, vendors. In the example in Figure 55, 90% compliance would save the organization an additional 4.5% of total spend on temporary labor.
The report in Figure 56 illustrates the amount of non-contract spend by each commodity area to allow an analyst to quickly focus in on the commodities and sub-commodities with the greatest potential savings opportunities, according to the total amount of non-contract spend and number of off-contract vendors being used.
After the analysis of off-contract spend has been completed, and all of the opportunities for savings by way of leveraged contracts have been addressed, the remaining spend should be re-examined against market indices wherever possible. Chances are that there are a number of categories where spending is significantly above index pricing and where there are additional opportunities for savings through smarter buying, even if volumes are smaller.

The next step is for analysts to identify these categories, determine pricing spreads, and create buying strategies for the buying teams. Then, the analysts need to define reports to track spending patterns over time to insure that the buyers are implementing the strategies and keeping spend in line with expectations[16].

Figure 57 shows the price paid for paper relative to the paper price index. The client had a contract in place that allows for price changes relative to the paper price index. Doing the detailed analysis showed that while the contract made sense, operationally price increases happen immediately on index increases, but price decreases lag behind index price decreases for a few weeks.
Without this comparison, the business unit would only see prices going up and down on a regular basis, which the business unit would believe to be in sync with the general industry based upon corresponding increases and decreases in the index. However, the business unit wouldn’t necessarily realize that price decreases lag while price increases are immediate. A contract that either fixed the price or mandated immediate changes upon index price changes beyond a threshold could deliver the organization significant savings.

Then Get Creative

At this point, the analysts will have used the spend visibility solution to identify the vast majority of opportunities that will deliver immediate payback as well as short term cost reductions. However, if the organization stops here, as many organizations do, it will see its return on investment drop off rapidly within six to eighteen months, in line with the classic value curve in Figure 1.

The reality is that these savings are just the tip of the iceberg. As per the example provided in a recent report on Strategic Spend Visibility, a properly designed strategic spend visibility initiative can deliver an ROI that is as much as five times greater than the ROI delivered by traditional, tactical, spend visibility efforts that stop once all of the low-hanging fruit have been picked.

Start with a Total Cost of Ownership

In any cost review, it is important to understand the total cost of ownership. Sometimes this is simple. But many times there are additional cost elements in the lifecycle, illustrated in Figure 58, which end up being material to the costs.
Sometimes, no matter how many suppliers are invited to the table, the price just won’t drop. That’s ok. If price is only one component of the lifetime total cost of ownership, focus on the other cost components of maintenance, warranty, and associated value-added service costs. The buyer might find that, in order to win the business, the supplier might be very willing to negotiate on these cost components. And, if not, there might be a third party willing to offer the organization substantially lower maintenance or service costs for a minimum commitment.

For example, if an organization is willing to give up the convenience of buying a PC and all related services from one vendor and instead buy the hardware, configuration services, and maintenance from different best-of-breed vendors, the organization could see a significant savings opportunity as illustrated by Figure 59.
Identify Underlying Cost Drivers

This involves going beyond a simple total cost of ownership calculation to understand the cost drivers inherent in the unit price. Let’s use the example of a company using a third party logistics provider to manage packing and delivery in a local geography.

For example, let’s assume that the vendor says that fuel prices are rising, and that he needs to increase prices by 20%. Does this make sense?

Let us further assume that a breakdown in the cost structure yields five main components:

- **Warehouse.** Cost to store the items getting ready for shipment
- **Boxes.** The cost for the cardboard boxes
- **Packing labor.** The cost to put the items in the boxes and prepare them for shipment
- **Vehicle costs.** The depreciation and maintenance costs for the delivery vehicles
- **Fuel.** The fuel costs for delivery.

The analysis, in Figure 60 then breaks down the costs into these five components and estimates the change in cost for each of the different elements.
<table>
<thead>
<tr>
<th>Component</th>
<th>Share of Costs</th>
<th>Cost Change</th>
<th>Percent Change</th>
<th>Resulting Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warehouse</td>
<td>20.0%</td>
<td>Flat, cost for heat</td>
<td>1%</td>
<td>20.2%</td>
</tr>
<tr>
<td>Boxes</td>
<td>10.0%</td>
<td>Cyclical</td>
<td>5%</td>
<td>10.5%</td>
</tr>
<tr>
<td>Packing Labor</td>
<td>20.0%</td>
<td>Flat</td>
<td>0</td>
<td>20.0%</td>
</tr>
<tr>
<td>Vehicle Costs</td>
<td>30.0%</td>
<td>Flat</td>
<td>0</td>
<td>30.0%</td>
</tr>
<tr>
<td>Truck Fuel</td>
<td>20.0%</td>
<td>rising</td>
<td>30%</td>
<td>26.0%</td>
</tr>
<tr>
<td>Total</td>
<td>100.0%</td>
<td></td>
<td></td>
<td>106.7%</td>
</tr>
</tbody>
</table>

**Figure 60 Cost Driver Analysis**

From this analysis, it shows that while a cost increase is warranted, it should be closer to 7% rather than 20%.

The team can now also use this information to determine where it might make sense to focus negotiations on price in the future. Supply Management might decide to lock in prices for the boxes or to lock in, or hedge, the organization’s fuel costs.

**Control Demand**

The largest cost savings on buying an item is not to buy it in the first place. Not buying something generates a 100% cost avoidance for each unit not purchased. Here are just a few questions that should be asked to generate ideas for controlling demand.

- What items or services is the organization buying?
- Which ones account for the most spending?
- Which ones are the most expensive?
- Is the organization asking the vendor for services which add little value but add significant cost?
- Is the organization buying based upon perceived value or proven value?
- Are there cheaper alternatives for the same purpose? Is there an easy process to automatically substitute cheaper items when more expensive ones are ordered?
- Does the organization need to maintain historical purchasing levels?
- Does the vendor have suggestions for cost reduction that have been ignored?
- Is the organization obsolescing items before they are used?
- Is the organization paying a premium because it prefers to do business with an incumbent vendor or a customer? If so, how much?
- Should the organization provide the service in-house or outsource?
- Can services be provided off-shore?
- Does another area (geography or business line) of the company already own the good or have service capacity which can fill a demand?
A good spend analysis effort always includes a demand management analysis. Adjusting what is ordered or the way services are delivered can save money without sacrificing quality.

A simple example is cleaning services. Many contracts have been written that all offices and retail space should be cleaned every day and all the trash emptied daily. For retail and high traffic areas, this can be important. For most offices, it’s not critical to have the carpets done every night. Perhaps every other night. Or even just once or twice a week. Pick a level that maintains the desired level of cleaning while generating the right savings, as illustrated in Figure 61.

<table>
<thead>
<tr>
<th>Full Service</th>
<th>Amount of Space</th>
<th>Frequency</th>
<th>Visits per Year</th>
<th>Cost per 1000 sq ft per visit</th>
<th>Cost per Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retail space</td>
<td>50,000</td>
<td>Most days including Weekends</td>
<td>350</td>
<td>10</td>
<td>$175,000</td>
</tr>
<tr>
<td>Office Space</td>
<td>50,000</td>
<td>All business days</td>
<td>250</td>
<td>10</td>
<td>$125,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total $300,000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Modified Service</th>
<th>Amount of Space</th>
<th>Frequency</th>
<th>Visits per Year</th>
<th>Cost per 1000 sq ft per visit</th>
<th>Cost per Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retail space</td>
<td>50,000</td>
<td>Most days including Weekends</td>
<td>350</td>
<td>10</td>
<td>$175,000</td>
</tr>
<tr>
<td>Office Space</td>
<td>50,000</td>
<td>twice a week</td>
<td>100</td>
<td>10</td>
<td>$50,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total $225,000</td>
</tr>
<tr>
<td></td>
<td>Savings</td>
<td></td>
<td></td>
<td></td>
<td>$75,000</td>
</tr>
<tr>
<td></td>
<td>Savings</td>
<td></td>
<td></td>
<td></td>
<td>25%</td>
</tr>
</tbody>
</table>

Figure 61 Cleaning Cost Analysis

Another common opportunity for cost savings is trash removal. Typically, the cost for removing trash in a dumpster is a function of two factors:

- **Hauling fee.** The cost to provide and move the dumpster; driven by frequency of the pickups.
- **Tipping fee.** The cost to ‘tip’ the dumpster into the landfill; driven by the local landfill costs.
Assuming the organization does not change the amount of trash it is generating, it can look at the size of the dumpster it uses. If it has a bunch of small retail locations with mostly non-food trash, it may be able to reduce its costs by increasing the size of the dumpster it uses and reducing the number of hauls.

For example, let’s assume the organization has a 2 cubic yard dumpster and that the hauling fee is fixed at $X$ for dumpsters between 1 and 8 cubic feet. If the tipping fee for a 2 cubic yard dumpster is $X$ and this is emptied every week, the total cost for a month is going to be:

**Two Cubic Yard picked up Every Week**

<table>
<thead>
<tr>
<th>Description</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pickups</td>
<td>4</td>
</tr>
<tr>
<td>Tipping Costs</td>
<td>$X$</td>
</tr>
<tr>
<td>Hauling Cost</td>
<td>$X$</td>
</tr>
<tr>
<td>Total</td>
<td>$8X$</td>
</tr>
</tbody>
</table>

If the organization replaces the 2 cubic yard dumpster with a 4 cubic yard dumpster, while the amount paid to tip will stay the same per cubic yard, the hauling fee will effectively be halved.

**Four Cubic Yard picked up Every Other Week**

<table>
<thead>
<tr>
<th>Description</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pickups</td>
<td>2</td>
</tr>
<tr>
<td>Tipping Costs</td>
<td>$2X$</td>
</tr>
<tr>
<td>Hauling Cost</td>
<td>$X$</td>
</tr>
<tr>
<td>Total</td>
<td>$6X$</td>
</tr>
</tbody>
</table>

**Rationalize the SKUs**

Some of the largest cost reductions are those associated with avoiding spend in the first place. This can often be accomplished by standardizing on parts and products (which eliminates entire spend buckets), improving inventory turns (which significantly decreases spoilage and obsolete inventory), adopting green solutions (which reduces waste disposal costs and negates the need for costly carbon credits in some geographic locales), and avoiding risky contracts and relationships in the first place. Each of these scenarios is possible when strategic spend visibility principles are brought to bear.
Not only will many large organizations buy many types of pen, paper, copiers, and other office supplies that can easily be standardized on a single component, but many organizations buy many types of components to include in their computers, cell phones, electronics, and consumer purchased goods when far fewer variants will suffice. While it’s true that a customer wants choice, a customer also wants value. Furthermore, a customer presented with too many options is likely to go into analysis-paralysis and either not buy at all, or buy from a competitor with a simple solution. When Ford decided that “... any color that he wants so long as it is black”, he was a visionary. SKU sprawl just adds complexity and cost, when sometimes it only takes one product to conquer the marketplace. (Just ask Apple.)

A good way to identify categories that should be subjected to a SKU reduction is to run a SKU utilization by category report and focus on those categories with a low total spend per SKU but a high number of SKUs. In the chart in Figure 62, the company is likely buying too many different types of routers, laser-jet printers, hard drives, memory modules, and flat panel displays. For an average company, there is no reason to buy dozens of different kinds of wired and wireless routers, more than one kind of B&W and color laser-jet printers per printing volume, more than a few hard drive sizes, more than a few types of memory configurations (desktop, laptop, and server) or more than two types of flat panel displays.

<table>
<thead>
<tr>
<th>Item</th>
<th>Quantity</th>
<th>Total</th>
<th>SKU / Category</th>
<th>Dollar per SKU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Servers</td>
<td>865</td>
<td>$3,605,046</td>
<td>18</td>
<td>$200,280</td>
</tr>
<tr>
<td>Notebook Computer</td>
<td>1,714</td>
<td>$3,184,523</td>
<td>4</td>
<td>$796,131</td>
</tr>
<tr>
<td>Workstation/Desktop Computers</td>
<td>2,971</td>
<td>$1,889,209</td>
<td>5</td>
<td>$377,842</td>
</tr>
<tr>
<td>Notebook</td>
<td>1,509</td>
<td>$820,276</td>
<td>4</td>
<td>$205,069</td>
</tr>
<tr>
<td>Flat panel displays</td>
<td>4,967</td>
<td>$773,872</td>
<td>13</td>
<td>$59,529</td>
</tr>
<tr>
<td>Memory modules</td>
<td>5,099</td>
<td>$520,375</td>
<td>15</td>
<td>$34,692</td>
</tr>
<tr>
<td>CPU Processors</td>
<td>378</td>
<td>$502,993</td>
<td>11</td>
<td>$45,727</td>
</tr>
<tr>
<td>Hard drives</td>
<td>1,351</td>
<td>$481,591</td>
<td>21</td>
<td>$22,933</td>
</tr>
<tr>
<td>Laserjet Printer</td>
<td>515</td>
<td>$277,901</td>
<td>17</td>
<td>$16,347</td>
</tr>
<tr>
<td>MULTIFUNCTION PRINTERS</td>
<td>579</td>
<td>$142,502</td>
<td>5</td>
<td>$28,500</td>
</tr>
<tr>
<td>Routers</td>
<td>2,835</td>
<td>$135,704</td>
<td>14</td>
<td>$9,693</td>
</tr>
<tr>
<td>Computer Screen Projector</td>
<td>118</td>
<td>$129,224</td>
<td>5</td>
<td>$25,845</td>
</tr>
</tbody>
</table>

Figure 62 SKU Utilization by Category

The analyst should do a quick review of the categories with a large number of SKUs to look for obvious cases where SKU reduction is possible. In the example summarized by Figure 62, if the company were to standardize on a high-performance blade platform capable of supporting virtual machines for each operating system run by the business, chances are that the server spend could be consolidated onto a single platform, which would represent a significant savings opportunity.
Analyze Inventory

An analysis of inventory costs will often indicate which items are staying in inventory longer than expected and where better forecasts and just-in-time production strategies can reduce not only associated inventory costs but the amount of working capital that is tied up in inventory. In addition, if the organization is using third party inventory management services, vendor managed inventory, and managing its own inventory, an analysis will reveal which option is the most cost effective for the organization.

A simple inventory turn analysis will quickly reveal the potential savings opportunity by speeding up inventory turns if the average inventory cost per day is known. The report in Figure 63 shows a typical breakdown of inventory turns for a consumer product goods company. If the company were able to convert significant amounts of monthly turnover to weekly turnover, it could free up over 20% of its inventory requirements in the best case (since 35% of inventory moving 75% faster frees up 26% of inventory).

<table>
<thead>
<tr>
<th>Turn Range</th>
<th>Days in Range</th>
<th>Inventory Value Snapshot</th>
<th>Number of Items</th>
<th>Percent of Annual Inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekly</td>
<td>1 to 10 days</td>
<td>10%</td>
<td>5%</td>
<td>43%</td>
</tr>
<tr>
<td>Monthly</td>
<td>10 to 40 days</td>
<td>35%</td>
<td>15%</td>
<td>35%</td>
</tr>
<tr>
<td>Bi-Monthly</td>
<td>40 to 70 days</td>
<td>30%</td>
<td>28%</td>
<td>15%</td>
</tr>
<tr>
<td>Quarterly</td>
<td>70 to 120 days</td>
<td>14%</td>
<td>12%</td>
<td>5%</td>
</tr>
<tr>
<td>Semi-Annual</td>
<td>120 to 200 days</td>
<td>6%</td>
<td>20%</td>
<td>1%</td>
</tr>
<tr>
<td>Annual</td>
<td>200 and over</td>
<td>9%</td>
<td>20%</td>
<td>1%</td>
</tr>
</tbody>
</table>

Figure 63 Inventory Turnover Analysis

Manage Energy

While the myth that green solutions are more expensive still pervades the marketplace, the reality is that a solution that is truly green will regularly save the organization green. For example, take IT. A green solution will reduce power consumption by 40% to 60% in the data center and 60% to 90% on the desktop. Considering that an average machine costs more to operate during its useful life than it costs to buy, this is very significant. In addition, any product that takes less energy or water to make or maintain will reduce overall lifecycle costs. Plus, any product or process that reduces carbon production not only decreases the (future) costs associated with carbon reduction or carbon credits, but increases the value of the brand.
Part of the challenge with energy management is that the costs are not visible to the regular user. The user plugs in the device to the wall – the energy costs eventually show up in some aggregated facilities cost. But they don’t see any direct value in their budget if they buy lower energy using devices. Procurement should work with business units to ensure they consider energy costs and how units can get recognition for reducing their energy usage. If it’s not in the RFP, it won’t be considered.

**Plan On Reporting As Part of the Process**

Every time you implement a savings initiative you should think about reporting on its success. Ask yourself “how will these results show up in the data?” and “How can I prove the benefits to a skeptical audience?”

If you negotiate new rates with a vendor, you should make certain that these new rates show up in the invoices. In the following example the invoice on the left provides only summary data. If you have just negotiated rates for master painters from $50 per hour to $45 per hour, you can’t see it. The invoice on the right shows the details and will allow you to both confirm that the new rates are in effect; and, you will be able to show people what the costs would have been before the new rates were put in place.

The additional detail in the invoice allows you to clearly report on the savings.
To make this reporting easy, you should arrange with your suppliers to provide you detailed reports on the detail from the invoices. This can easily be done in a simple spreadsheet delivered along with the invoice. This data should be easy to load into your analysis system and, combined with historical pricing, should be able to provide you with concrete savings reports.

Every vendor for every commodity can provide you with this level of data. If the vendor can’t provide item level pricing information, you have to ask how they are producing their own invoices. Obviously they have the data, or they could not have invoiced in the first place.

We have discovered that this reporting need will impact how you contract for items. Many contracts have very complex and convoluted pricing structures. If the pricing is hard to understand, it’s likely hard to verify. A good pricing structure in a contract should support quick analysis. You may ask for a great deal of information in your bidding process, but the final pricing should be as simple as practical, without giving up on the new pricing you have negotiated.

**To Each Her Own**

A spend visibility effort is only truly successful when it pervades the organization and allows everyone who needs access to up-to-date spending information to access that information when, and how, they need it. To this end, the platform will need to provide different capabilities and views for each of the primary users of the system.
C-Suite Executives

The CEO, CFO, COO, and other chief officers need quick insights into current organizational spending, current projects, and (immediate) problem areas. They are primarily interested in reports. In essence, what a CXO wants is a set of reports that summarizes what she needs to know, when she needs to know it, in a manner that is easy for her to comprehend. She doesn’t want a dashboard that never appears to change and that doesn’t give her useful information.

This being said, it is a rare Senior Executive in a large company who logs directly into a system to monitor what’s going on. They have someone prepare a report to show them what is important. The person preparing these reports needs a way to quickly and accurately create these reports.

CPO and Procurement Managers

The CPO needs the ability to quickly get information on dozens of different tasks and metrics to track the current status of the Procurement organization. For example, at any given time the CPO might need to know:

- Total Spend
- Number of vendors
- Spend by commodity
- Channel use: spend, transactions, bypass
- Preferred vendor use: spend, transactions, bypass
- Contract status
- Project status

All of this information cannot be summarized on a single screen. The CPO will need a series of reports or a system that provides high level information at a glance but lets her quickly get to the information of relevance.

Different metrics will require different views, and the best way to illustrate this is with a single metric. Consider the “Use of Contracted Vendors” metric, where the CPO will want to know the overall spend, the breakdown by business unit or commodity, and bypass spend. The CPO will need to start with a view similar to the one presented in Figure 64.
After perusing this table, the CPO, who probably already knows that emergency projects resulted in 50% of the operations and headquarters bypass spend, would probably want to know why Retail, who experienced no major changes in operations, has a bypass spend of 27%. At this point, the CPO will need to dive into the Retail unit to try and find the source of the bypass spend. The CPO will need to produce a view similar to the one presented in Figure 65.
At this point the CPO will notice that the worst offenders are the Consumer and Small Business, Consumer Lending, and Small Business Banking groups. Furthermore, the CPO will likely suspect that due to the similarities in focus (Consumer), that these groups are likely using some of the same vendors and that a quicker way to track down a significant portion of the bypass spend might be to ask which vendors are being used by the retail groups. At this point, the CPO will want to produce a crosstab that plots vendors against retail groups, as in Figure 66.
At this point the CPO will have confirmed some of her suspicions, since the biggest offenders are Chex Systems, Deluxe Check Printer, Market Intelligence, and Credit Data of California, and all but one of these are used by multiple retail groups with large amounts of bypass spend. At this point, the CPO will likely want to drill into specific vendors for a more detailed breakdown by cost center, as in Figure 67.

She may want to then drill down into the individual Invoice Payments for the vendor and business unit intersection(s) of interest, as in Figure 68.
Most of the time it will not be the CPO or Procurement Manager doing this analysis. Usually, an executive or manager will raise the question and the analyst will do the rest of the investigation. It is extremely useful to be able to drill down in a single path to gain the understanding.

**Analysts**

Analysts need detailed volume, cost, and utilization information before they strike a contract. They need to know what the major contributors to cost are in order to focus their negotiations for maximum impact. They need to understand what each proposed concession really means and whether a deal on the table is one they should accept, counter, or run away from. They are the power users using the spend data to drive the next level of savings initiatives who need the full suite of capabilities available in a modern spend visibility and analysis solution.

The tables in Figure 69 and Figure 70 illustrate the type of data an analyst would like to have before starting a negotiation for medical supplies for a hospital or medical supplies GPO. In the example in Figure 69, the data is limited to “gloves” for simplicity. At a high level, the analyst will want to see volume, average, and total cost information.

![Figure 69 "Glove" Report](image-url)

Then the analyst will want to dive into each type of glove in the commodity family and see the different types (sizes and medical specifications) being purchased.
With this data, the analyst sees that the greatest opportunity for costs savings are in the Size 7 class. Now the analyst needs to see historical pricing to get a feel for the commodity. The best way to do this is through a bubble-chart pricing trend analysis, illustrated in Figure 71, that plots the average price paid over time.

Looking at this data, the analyst sees that there has been significant price variation over time, with several instances of multiple prices being charged for the same product in the same month, and that the price has steadily increased over time. This tells the analyst that there is an opportunity for cost savings on this category if a good price can be locked in, or at least tied to a plastics index.
10 Refresh, Repeat

The next step to improve the analysis is to refresh the spend data. The data needs to be kept current. This should be done at least monthly, and perhaps weekly for fast moving inventory. (Generally, data does not need to be refreshed more often than weekly because there is no “real time” in AP spend analysis. By the time a transaction is in the system, the money has been spent.) This needs to be done for two reasons. One, as new spend categories are introduced, new savings opportunities will present themselves. Two, many savings opportunities are dependent on the buyers following prescribed savings strategies. As a result, the analysts need to be able to track whether or not the strategies – such as supply base consolidation, maverick spend reduction, or SKU rationalization – are being implemented.

The emphasis on refresh and tracking, especially for spend categories where spending decisions have been recently made, should be on compliance – with spending decisions, organizational policy, and external regulations that impact the organization with respect to labor, trade, and the environment.

The system providing the transaction data should be able to provide the incremental data in the same format as last month’s (incremental) data. The system used for spend analysis should automatically be able to apply all the logic to the new month of data and provide a set of reports that answers “what has changed”. The analysts just need to review exceptions and items not dealt with in the existing rules.

The high level process is:
1. Add new transaction files, check contract totals
2. Add updated index files, thin index
3. Vendor grouping for new vendors
4. Commodity mapping for spending not mapped
5. Quality control
6. Distribution of the data to users

This process should be completed in a fraction of the time it took to create the original dataset.

Many companies refresh their data on a monthly basis. As this data is refreshed, reports are sent out to users to help them understand what is happening with their spending. Different users will be looking for different attributes.

The graph in Figure 72 shows the monthly spend for the top level categories in a typical company hierarchy. The reader can see how their spending is trending over time and how their spend is changing relative to other categories.
This report in Figure 72 could also be done at a business unit level, as per Figure 73.

When the analyst sees spend by business unit over time, they will be able to quickly identify which business units are increasing their spend. In this case, the analyst might notice that HR spending has increased significantly over the last few months and the facilities has steadily risen over the last 12 months. Both of these trends should be explored to understand if they are expected increases or if something unexpected is occurring.
In a vendor-count time series report, such as the one in Figure 74, the analyst can see the year-over-year trends and the number of new vendors that were added in the last month. If there is an initiative to reduce the number of vendors, it is easy to quickly identify which areas need attention and which areas have been complying with the initiative.

![Vendor Count by Commodity Time Series](image)

**Figure 74** Vendor Count by Commodity Time Series
Finally, when all is said and done it is time to go back to the beginning and start all over again once all of the initial analyses, which could take quite a while for a large organization, have been completed. Over time, new opportunities will appear in each of the categories. As the organization adapts to the current strategies used by less-than scrupulous commodity suppliers to bill more than they should, these suppliers will either make new billing mistakes, or find new ways to overbill over time, and new opportunities for cost recovery will present themselves. As new products gain dominance in the market place, the potential supply base will expand from just a few suppliers with high prices to a large number of suppliers with extremely competitive prices. These categories will once again become low-hanging-fruit opportunities. As new product offerings are introduced by the organization in an attempt to penetrate a new market, SKUs in new categories will begin to proliferate throughout the organization. Thus, once the organization has gone through the full cycle of strategic spend analysis enabled by its strategic spend visibility initiative, there will be new opportunities for cost recovery, new low-hanging fruit to pick, and new opportunities for strategic cost reduction through supply base consolidation, SKU rationalization, and fraud reduction.

**Best Practices**

While it's true that an organization can buy a visibility solution, roll up their sleeves, grease their elbows, dive in, and, with a lot of hard work, obtain some decent results, the organization will get better results faster if it applies some tried and true best-practices. This section reviews some of the best practices that have proven to be the most successful in spend visibility.

**Formulate a Plan**

The first step is to formulate a plan of attack. Start with the basic process defined in the last section and customize it to specific organizational needs. Specifically, focus on identifying suspected capabilities and gaps that should be explored, specific cleansing and classification rules unique to the data organizational schemes in existing enterprise systems, the starting cubes and reports, and the first set of analysis targets. This will not only insure that the visibility initiative runs smoothly, but that the analysts are able to get to work as soon as 80% of the data is amalgamated and mapped.

**Analyze Everything**

Traditionally, the classical decision with regard to low probability ad-hoc data analysis has been not to do it based on the fact that, before the advent of modern spend visibility solutions, it was time consuming and very costly in terms of human resources.
However, with a modern spend visibility and analysis system, an analyst can build a rich cube (with base and derived dimensions) of millions of transactions on her high-end multi-core laptop in a matter of hours and analyze it in real time. Whereas the cost of low-probability ad-hoc data analysis used to be in the thousands, if not tens of thousands of dollars, the cost of a low-probability ad-hoc analysis is now only a few hundred dollars. As a result, the decision should be to do it and follow any hunches the analyst might have. After all, if the probability of success of these hunches averages 1 in 10, then, for every 10 hunches an analyst follows, on average, 1 will pay off. It only takes a couple of savings successes for this new-age spend visibility strategy to pay off handsomely.

When it comes to spend analysis, it’s important to remember what a VP of a large insurance company said in Aberdeen’s Spend Analysis: Working Too Hard for the Money[7]: “If you take away spend analysis capability from a company, you might as well turn the lights out on the procurement department. Then all they’ve become is a transactional processing group that doesn’t add much value to the organization.”

**Build as Many Cubes (and Reports) as Possible**

The key to spotting savings opportunities is to look at data in as many ways as possible because it’s impossible to know up front which cube (or report) will yield the most productive insight. As highlighted in the recent white paper on Strategic Spend Visibility[6], spend analysis is much more than just one AP level spend cube. It’s many cubes on many different types of data, including multiple cubes by supplier by commodity for contract compliance. It’s cubes for throw-away one-off analysis, cubes that are derived from other cubes, and so-on. It’s cube after cube until the analyst finds the one cube that gives her the insights she requires to help the organization realize massive savings.

For example, a simple breakdown of total spend on cellular communications by department might indicate that the department breakdown is in line with expectations with Sales spending the most, followed by IT and then Management with total costs in line with average spending for a similar sized company. However, unless the analyst looks at average costs per minute, the analyst would never know that Sales is spending more per minute than industry average due to a bad contract selection and that IT is actually spending less per minute than industry average and that their contracts should not be altered.

To see why multiple views, and cubes, are often needed, consider the first-pass analysis of telephone usage by business unit in Figure 75 which presents no obvious outliers and no obvious savings opportunities.
If the analyst perseveres, and does a detailed analysis of the top users by cost and minute, as in Figure 76, she will quickly see that there are a number of users whose plan (500 minutes) does not cover their usage. It will also reveal a class of users who theoretically have unlimited plans but still have a high monthly cost. These plans will have to be investigated to determine if there are billing problems that need to be resolved or if the user is doing a lot of roaming (internationally) and add-on plans are required to keep costs down.

<table>
<thead>
<tr>
<th>Last Name</th>
<th>Minutes</th>
<th>Cost per month</th>
<th>Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>HARDEL</td>
<td>1445</td>
<td>$1,395</td>
<td>500 Min</td>
</tr>
<tr>
<td>BIELECKI</td>
<td>1348</td>
<td>$1,298</td>
<td>500 Min</td>
</tr>
<tr>
<td>PASQUARELLA</td>
<td>1157</td>
<td>$1,107</td>
<td>500 Min</td>
</tr>
<tr>
<td>TERADA</td>
<td>1085</td>
<td>$1,035</td>
<td>500 Min</td>
</tr>
<tr>
<td>SMITH</td>
<td>803</td>
<td>$753</td>
<td>500 Min</td>
</tr>
<tr>
<td>DIGIACOMO</td>
<td>725</td>
<td>$675</td>
<td>500 Min</td>
</tr>
<tr>
<td>FARON</td>
<td>514</td>
<td>$464</td>
<td>500 Min</td>
</tr>
<tr>
<td>GUILL</td>
<td>1480</td>
<td>$1,200</td>
<td>Unlimited</td>
</tr>
<tr>
<td>HEAPS</td>
<td>1338</td>
<td>$1,143</td>
<td>Unlimited</td>
</tr>
<tr>
<td>BEDOYA</td>
<td>1264</td>
<td>$1,045</td>
<td>Unlimited</td>
</tr>
<tr>
<td>BISH</td>
<td>1099</td>
<td>$1,025</td>
<td>Unlimited</td>
</tr>
<tr>
<td>FRY</td>
<td>899</td>
<td>$967</td>
<td>Unlimited</td>
</tr>
<tr>
<td>DAVISON</td>
<td>891</td>
<td>$75</td>
<td>Unlimited</td>
</tr>
<tr>
<td>HIGA</td>
<td>878</td>
<td>$75</td>
<td>Unlimited</td>
</tr>
<tr>
<td>KACZOR</td>
<td>576</td>
<td>$75</td>
<td>Unlimited</td>
</tr>
</tbody>
</table>
In addition, if the analyst also looks at the users who use the fewest minutes, as in Figure 77, she will likely find a number of users who aren’t using their phones at all. These phones should be considered for elimination, or, if they are needed for emergency (such as on-call), replaced with a small number of “business unit” phones that are lent out on an as-needed basis.

<table>
<thead>
<tr>
<th>Last Name</th>
<th>Minutes</th>
<th>Cost per month</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAULDEN</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>STUCKY</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>MATHEWS</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>WATTERS</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>MASSIE</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>MATAMOROS</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>TILLEY</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>KOHOUT</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>SLOVAK</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>MOGENSEN</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>CHRISTIAN</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>NUCKEY</td>
<td>0</td>
<td>$50</td>
</tr>
<tr>
<td>ABBOT</td>
<td>0</td>
<td>$50</td>
</tr>
</tbody>
</table>

*Figure 77 Low-Usage Cellular Analysis by User*

Furthermore, be sure to create a number of cubes or reports that use range dimensions on both chronological, spend, and classification dimensions. Sometimes the best insights come from looking at chronological spending patterns month-over-month, quarter-over-quarter, and year-over year. Sometimes the best supply base consolidation opportunities come from segmenting suppliers into small, medium, and large spending buckets and noticing which bucket has too few or too many suppliers. And sometimes the best cost avoidance opportunity comes from analyzing the breakdown of spend between low, medium, and high risk suppliers and analyzing the savings offered against the risk of loss.

**Drill, Drill, Drill**

A savings opportunity is not always what it appears to be. For example, a year-over-year comparison of the 20M telecommunications spend might reveal that spend is up 2M or 10%. If the organization hasn’t considerably increased its headcount and the market has been competitive, then this would, on the surface, indicate that there is likely a significant savings opportunity to be had in this category. However, it might not be through renegotiation, which could be the analyst’s first instinct.
For example, if the previous year saw the execution of an initiative to reduce costs by 10% (under the previous analyst) which resulted in new contracts, then year-over-year spend should be less. While the full 10% wouldn’t be realized if usage by existing staff shot way up (which isn’t likely unless Procurement replaced all their cell phones with data hungry 3G smart phones and didn’t choose a data plan wisely), there should still be a noticeable and significant drop in spending. At this point the analyst might conjecture that the providers are not billing at the contracted rate in some cases, but is this the case? And if so, where are those cases? And is it a worthwhile effort comparing long multi-line invoices to convoluted pricing contracts when that might not be the problem at all?

The answer is that it’s not worth a line-by-line comparison of each invoice against contracted rates, unless an analysis indicates that is the problem, especially when five or ten minutes of intelligent drilling can quickly identify where the problem is and where to look. For example, let’s say that the end result of the savings initiative required all divisions and business units to switch long distance carriers and buy off of a new master contract. If all of the divisions and business units made the switch, then the instinct of overbilling might be right. But if some of the larger units did not switch, that could easily account for the overspending – which can easily be detected with some simple drilling. If the analyst simply drills down the hierarchy, fixing the divisions and units where spending is up at each step, it will only take a few drills until she finds the thirty (30) intersections of division and business unit, out of the hundreds of divisions and units across the organization, that are spending considerably more than last year. A quick drill into the transactions of each unit will then quickly reveal that fifteen (15) of these didn’t switch carriers and are still paying twice as much, on average, as they would on the new contract. A simple report to the CFO and the head of each unit then corrects the problem and the negotiated savings will then start to quickly materialize.

Another example is office supplies. Overcharging is common in this category. A quick drill down into a cube with invoice and contract data can quickly determine how much is being overcharged by product, and who approved the overcharges. The table in Figure 78 identifies over $54,000 in overcharges in five categories, with 70% of the overcharges being approved by three individuals. A quick investigation into the purchasing of these three individuals will likely result in savings of over $30,000 on these products alone. In this case it was found that the individuals were ordering from the wrong catalog.
Sometimes a supplier might think they are giving the organization a great deal, and be unwilling to come down on price, until they see the total cost of the relationship. Once they see a report that clearly aggregates and break down the unit costs, shipping costs, storage costs, tariffs, and utilization costs – including warranty and return costs, as well as the associated service and support costs – for all major categories, they might realize that what is a great deal from their perspective is not such a great deal from the organization’s perspective. At this point, they might be more likely to work with the organization to identify process and product improvements that will save both parties going forward.

**Increase Spend Under Management**

The ultimate key to savings success in any organization is to increase spend under management. This is done by analyzing spend in order to build up an ongoing opportunity pipeline. The goal of every analysis project should be to produce a spend management plan for the commodity, category, supplier, geography, or organizational unit under analysis. This plan may involve a contract, a supply base or SKU rationalization, a compliance effort, or simply maintenance of the status quo if the current strategy is reasonable.

As long as the strategy is documented and the results are tracked, the organization is on the right track. Increasing spend under management is the only true way to see long-term results from a spend visibility initiative.
Treat All Categories as Equal

Some organizations believe that spend analysis is useful only for direct spend categories, because that’s where the majority of spending is. While that may be true in traditional manufacturing industries, it’s not true across the board. Typically, indirect costs for office workers are equal to their salary expense. If COGS is 50% of your expenses, the other half is probably roughly split between salaries and vendor expenses. This means that the indirect spend is a significant portion of the spending. In some industries, like health care and call center management, services spending can be as high as 80% to 85% of the total vendor spending. When it comes to spend visibility, no category should be overlooked. One never knows where the greatest savings opportunities lie. There might be an opportunity to decrease purchase costs in a component category by buying a raw material, like steel, at a higher discount on behalf of the component suppliers, or there might be an opportunity to reduce temp labor costs by optimizing contract distribution between national and local placement agencies. There is no way to know until the analysis is done.

Share

The best results come from a collaborative effort where the data is shared not only between the analysts, but between the buyers and end-users in the various parts of the organization who can then understand the impact of their decisions on spending. By sharing the custom cubes and reports that she develops during an analysis effort, an analyst increases the chances that she will be notified of a potential savings opportunity that could be revealed by a related analysis she did not think of.

Work with others to establish the value of analytics. Saying “we did a great job last year” is much less credible than “Finance reports that our actual savings figures are $5.2 million for the current year” or having a business say “the purchasing analytic team helped us generate a $1.2 million savings”.

Create a Cube-Based Knowledge Store

Any time an analyst produces a cube or report that increases insight into a category or savings opportunity that results in a savings strategy, the cube or report should be stored in a centralized knowledge store. Not only will this simplify tracking of the results of the savings initiative and compliance with the savings strategy, but it will provide a foundation for future analysis efforts (as it will simplify the year-over-year spending analysis) and insure that the results are sustainable. Furthermore, a cube-based knowledge store provides a great foundation for training of future analysts.
Measure and Report the Results

The best way to get recognition and support for future analysis and projects is to measure the success of each and every project and report the results regularly.

With respect to reporting, a summary could be written for the monthly electronic newsletter. In addition to expressing the results in hard dollar savings, this summary should also link the savings to the corporate financial metrics that will get the attention of the C-Suite, including return on investment (capital) or ROI(C), weighted average cost of capital or WACC, and economic value add (EVA). Once the C-suite sees the positive impact that spend analysis is having on the balance sheet, it will be easy for the analysis team to get the hardware, resources, and support that they need for more analysis capability.

Train, Train, Train

A great analyst is dedicated, creative, and well-educated. A great analyst is up-to-date on modern analysis techniques, cost reduction strategies, and innovation strategies that will allow the organization to find new ways to consolidate the supply base, rationalize its SKUs, reduce its carbon footprint, and effectively reduce the lifetime total cost of ownership of every product and service it purchases. And a great analyst can import data, build a spend cube, map data using rules and overlays, create ranged and/or rolled-up dimensions, create reports, and drill into multiple dimensions in a matter of hours and then repeat the entire exercise to test a different hypothesis after lunch. Until the senior analysts are at this level of proficiency, they need training. Each analyst will need two to three weeks of training over the course of a year as they work their way up from beginner to expert.

Continue to provide advanced training and knowledge sharing so that people don’t become stale and they continue to expand what they know how to do. You might even do a quarterly “guess what we did with data” session.

Actively Look for Strategic Opportunities

With so much noise in the marketplace about AP analysis, which is tactically oriented, it’s easy to forget that the tactical savings opportunities identified from an analysis of AP data are only the tip of the iceberg. True success is the result of a strategic spend visibility effort that goes beyond AP and even invoice data to analyze all organizational spending in a focused and methodical manner that will uncover each of the savings opportunities outlined in Strategic Spend Visibility: Untapped Potential for Cost Reduction[6] and then some.
With the right spend visibility, spend analysis is a powerful tool, especially considering that it enables the average organization to identify a savings opportunity of 11% ... much more than the average organization could expect to save with a simple e-Negotiation, Contract Management, or Supply Management solution. In fact, there's only one other strategic sourcing tool that provides an organization with such a significant savings opportunity, and that tool is strategic sourcing decision optimization. Two back-to-back studies by Aberdeen in 2005 and 2007 found that an average organization could expect to save 12% more through decision optimization than it would if it just used standard e-Negotiation tools like e-RFX or e-Auction.

Decision optimization complements spend analysis quite nicely. One of the major drawbacks of optimization is that, even today, if the model is big enough and complex enough, it can still take hours on even the highest end server to find an optimal solution. While this is significantly better than the days, or weeks, that a sourcing team used to have to wait for an analysis to be conducted, it's still a delay and an interruption to the sourcing process, because it means queuing up a bunch of what-if scenarios overnight and then coming back the next morning only to repeat the exercise if none of the runs resulted in an acceptable award scenario.

However, when analytics is applied to cost data, it's very easy to build cubes, crosstabs, tree-maps, and reports that rank the lowest cost suppliers, the highest quality suppliers, the lowest cost suppliers relative to a particular quality metric, or any other organizational requirement for supplier or product selection. Now it's easy to identify the suppliers who don't make the cut in quality or on-time delivery metrics, the suppliers who can't handle the capacity, or the suppliers who are just too costly to be considered, as well as any products supplied by the remaining suppliers that can be eliminated. In a large category, the end result of an exercise like this would typically be a model that contained only a dozen suppliers, a hundred products, and a few thousand lanes instead of a model that contained dozens of suppliers, thousands of products, and tens of thousands of lanes. In other words, the result is a model that is an order of magnitude smaller with a complexity that is orders of magnitudes less. This model would likely solve in minutes instead of hours and allow the analyst to continue the analysis process uninterrupted; which means she could build what-if scenario after what-if scenario which can be successively fine-tuned until a plan to capture the savings is realized.
Don’t Be Afraid to Outsource Analysis

Sometimes the only way to find the savings is to outsource to a professional spend analyst who not only has senior data analysis skills, but who also has the deep category expertise required to identify which particular patterns of spending represent true savings opportunities. For example, there are consultancies that specialize solely in the analysis of telecom and utility spend. While it may be obvious to the senior analysts that telecom represents a significant savings opportunity, it may not be obvious what the best strategy is to realize that savings, or even how much of a savings opportunity there is. However, an expert consultant will quickly be able to identify the opportunity, the strategy to realize the savings, and assist in the implementation. Most organizations are resource constrained. Do a cost benefit / timing analysis to understand how long it will take you to pursue the opportunities if you do it all yourself.

Selecting the Right Platform

There are a number of platforms on the market, and while some are arguably better than others, the most important part of platform selection is that the platform provides the foundations necessary to implement a strategic spend visibility effort that will meet organizational needs. While the specifics might differ from one organization to another, the following eight requirements should be considered universal and are generally necessary for any spend visibility effort to succeed.

Great (E)TL Support

Simply put, kick-ass (Extract-)Transform-Load capability is critical as the system will have to import and process multiple data feeds coming from multiple sources in multiple formats and normalize all of the data into a common schema. Furthermore, any user must be able to accomplish this task quickly and easily whenever she needs updated spend information or needs to explore a hunch in an ad-hoc analysis.

Much ado is made of Extract-Transform-Load, or ETL. There is a need to get data from the source system into the analysis tool, but it may not require a separate ETL system. A good spend analysis tool will often be able to do what’s needed.

The data extractions allowed will depend on the source system and the data access controls in the data center. Most data centers can “run this query monthly”. That will likely get the data out of the source system. The data center / IT team will usually be resistant to adding new software to run extracts and it will take months to get this done. Most large companies already have solutions for this problem. Some companies have specialized query tools already, some use SharePoint to transport data. Avoid the problem by using the existing data extract method and don’t reinvent the wheel. And don’t try to tell the IT group how to do their job.
The transform and modifying of data is a key part of the analysis process. It will be done hundreds or thousands of times. The process is never “one and done” but rather is a voyage of exploration. The transformation of the data has to be part of the spend analysis system, not the data loading process. If the spend analysis system doesn’t allow you to regularly transform the data, then the organization should review what the system does for it.

**Real-Time Changes**

The spend visibility solution must support a reasonable level of spend analysis in real-time. A user must be able to not only rapidly drill down into existing spend cubes, generate reports on selected data and dimensions, and export selected spend to pivot tables in Excel but be able to define new cubes, views, and reports on reasonable sized data sets (of at least 5 - 10M transactions) and have the cubes, views, and reports available for analysis in a few minutes. Four hours to rebuild a cube is simply unacceptable when there are solutions that can rebuild a cube of up to 10M transactions on a laptop in under 5 minutes.

Now, if the analyst is trying to rebuild a cube of 100M to 1B transactions, then that’s a different story. However, in this case, it is arguable as to how much value can be extracted from a cube this big. The best results usually come from focused analysis of suppliers, categories, geographies, or other data sets with highly correlated data. Few insights generally come from trying to analyze all of global spend simultaneously ... it’s analogous to trying to compare the ROI of the organization’s janitorial service investment with the ROI of its Big 5 consultancy investment with the ROI of its investments in the stock markets. Does that really make sense?

**Intuitive**

The platform must be so easy to use that the average user can pick up basic usage quickly without training. This is critical if the organization is going to achieve the widespread adoption by analysts, buyers, managers, and CXOs that is necessary for the spend visibility solution to truly be considered a success.

This does not mean that the platform has to use FLASH, HTML5, AJAX, or some other modern web-based all-flash-and-no-substance consumer user-interface (UI), but that the UI has to be clean, concise, logical, and laid out in a manner that makes sense to the user and allows her to accomplish what she needs to quickly and easily. The process to get from task A to task B should be simple, clear, and easy to follow. The user should not have to go through screens W through Z on the way to B from A.
**Visual**

While 3D graphics or 1 billion colors are not a requirement, the package must be capable of rendering decent graphs in multiple colors that contrast well, since most people will grasp a visual display more quickly than a large table with hundreds or thousands of summary rows. The user must be able to define common Excel-style charts and graphs on any set of data and render them easily.

**Resource Light**

The spend visibility solution should not require a lot of resources to install or maintain and should be entirely self-service in the hands of the analyst group. The analysts should be able to import the data that is required, analyze it, and export any results that need to be saved. There should not be a lot of support needed from IT – preferably none – or from the vendor, as this will raise costs and can impede implementation, analysis, and, most critically, adoption.

**Competitive Cost**

A good solution that meets all of the requirements outlined in this spend visibility implementation guide should start in the five figure range for an average mid-sized organization. Long gone are the days were it cost a million dollars just for the data warehouse that was required for the overpriced BI (Business Intelligence) tools. The organization should understand all the costs, including initial data loading, training, software licensing / SaaS (Software-as-a-Service) costs, hardware required, and maintenance. If the software is sold traditionally, i.e. like Microsoft Word, where the product must be periodically re-purchased in order to "stay current," that cost must also be factored in.

**Output to Formatted Reports**

The executive team will be very interested in the data and they will want the data included in reports and presentations. The system has to make this easy and repeatable. The output needs to be more than transaction dumps. For example, can the user output some summary data, a pivot table and some graphs all into a spreadsheet with a single mouse-click?

**Performance**

The system needs to provide very fast response to queries. If users are waiting minutes for responses, they will use the system less than they should.
Be sure to ask whether or not users can create new dimensions or mapping rules all on their own. Find out if they can create a custom range dimension or create a new hierarchical mapping of spending in a new cube, for example dividing it into two groups to understand the cost center implications. If there is a limit to what analysis can be done, and what data transformations can performed by the end user, then there are serious limitations on the results that can be obtained.

**Successful Test Drive**

Before the organization invests in a Spend Analysis system, it should see its data in the system. It doesn’t have to be all of the organizational data, but enough to yield a sensible analysis and let the analysis team see the system in action. If the system is hosted by the vendor (e.g. SaaS), they should be able to make an instance temporarily available. If the system runs on internal hardware, the spend analysis group should get a machine and test the software for a few days or weeks. If the vendor refuses to do this or charges a large amount for this, this is a good foreshadowing of what making changes will be like in the future.

The Test Drive should be long enough to ensure that a spend analysis can accomplish the basic functions of:

- **Data load** – using the built in (E)TL tool,
- **Data transformation** – via user-defined mapping rules,
- **Data indexing** – on user defined attributes and dimensions,
- **Data analysis** – cubing, slicing, drilling, cross-tabbing, dimensioning, and reporting, and
- **Data update** – with the next month / quarter / year of information.

This is accomplished by having a vendor representative walk a senior analyst through each stop of the process using the organization’s data. Depending on the organization’s size, start by exporting all of the transactions for a quarter, year, or three year period in to a flat-file. This file should, preferably, contain at least 100,000 transactions, but no more than ten million. This initial transaction data set needs to be large enough to insure that the tool can handle a high volume of data, but not so large that it would be unreasonable to expect the data set to be processed in real time on an average laptop. (Leading products can process over ten million transactions on a moderately high-end dual-core laptop with 4GB of memory in a few minutes. As a result, a multi-tenant solution on a multi-core SaaS machine should be at least as efficient, if not more so [depending on user load and internet latency].)
In addition, the test drive should demonstrate any and all claims made by the vendor with respect to performance, mapping, etc. on the organization’s data. If the vendor is making grand claims about its tools, make them prove their claims, on the organization’s data, in real time.

If the vendor passes the initial sniff test, work through the following steps with the vendor representative to make sure that the tool is usable by the organization:

- **load the data**
  this should be easy and only take a few minutes.

- **examine the raw data**
  it should be easy to spot top vendors, cost centers, and GL codes.

- **group some vendors**
  this should be easy and immediately reflected in the top N reports.

- **create a high level commodity structure and map some spending**
  use the “secret sauce” of map the vendor, map the GL codes, map the vendor + GL codes for the top n vendors and GL codes and drill around the cube; it should be easy to see the spending break down at a high level.

- **modify some rules and create a few exception rules**
  the reports should change in an obvious way; drill around the cube and find the changes.

- **create a custom report**
  print it out, compare it to known spending patterns, make some mapping changes, run the custom report again, and compare it to the first version.

- **alter the report**
  make sure it’s quick and easy to do, because no canned report survives first contact with a manager.

- **perform an update and refresh**
  use additional data from previous or succeeding quarters or years to update and refresh the cube.

This process should not take more than half a day to a day. If it takes more than seconds to make a real-time update to the cube when grouping vendors or mapping spending, you need to consider whether the tool is sufficiently flexible. If it’s a big deal to change things, then analysts won’t bother to do so; and important opportunities for savings will be missed.

As soon as the analysts have verified that the tool can do what is needed, negotiations to lease or own should begin immediately. Every day that analysis is postponed is another day during which the organization is spending more than it should.
Selecting the Right Architecture

Does the organization want, or require, an on-premises solution? Is the organization leaning towards a remotely hosted, SaaS (Software-as-a-Service) thin client solution? Or does it want a hybrid solution where the application is run on-site but with a back-up in the vendor’s data center for data security? Or is an installed (fat client) solution a better alternative?

If the organization believes that it needs an on-site or behind-the-firewall solution because of data security concerns, reliability, or the need for tight integration into existing systems, that conversation should be the first conversation with any potential vendor.

Does use of the system require any IT skills or "expert" support? Some solutions run right on conventional hardware, and require no IT support at all. Others, such as conventional Business Intelligence systems, require a core expert user community.

If the solution does require IT support, are the resources and skills available? If the spend analysis team can’t run the solution, then the organization will need someone else do it, either on their hardware and on their premises with their resources, or on a remote server and with remote resources that the organization does not control.

The best approach for an organization is generally one that fits with the organizational philosophy on IT systems and offers the path of least resistance. Each has its own advantages and disadvantages.

If the organization does not have a preference, and is highly cost sensitive, it should look seriously at SaaS or fat client (installed) solutions that can run stand-alone on the desktop, especially if the organization does not have IT resources to spare. If the platform was built properly, SaaS and fat client solutions can often provide a lower-cost alternative.
Summary

Understand spend analysis principles

Spend Analysis is not a one-time exercise, but rather a continual process. It is not restricted to AP data and includes all organizational data related to spend, service, and even quality. It is not a set of canned static reports but an ever-changing view into your company’s spend. It is not a set of rigidly defined analyses repeated at regular intervals, but a set of ad-hoc initiatives that follows the money and the gut-feel of your trained and talented analysts. With a modern spend analysis tool, an analyst can follow her instinct and analyze a new subset of data in a matter of hours. Even if the analyst makes $100 an hour, each analysis takes 4 hours and only one in ten analyses pays off, the effort will return you 10X if she finds a 40,000 savings opportunity and 100X if she finds a 400,000 savings opportunity.

Identify cost and demand drivers

True spend analysis doesn’t stop by identifying the top ten savings opportunities based upon volume, current spend, and current market pricing. It dives in and tries to understand why costs are rising or staying high and why the organization buys what it buys. By doing so, the organization is not only able to get a better grip on what costs should be if key component costs (such as energy, petroleum, wheat, or steel) rise but also able to negotiate reasonable year-over-year increases. Transportation is a prime example where understanding the impact of a fuel increase can save your organization a lot of money. If fuel goes up 10%, that doesn’t mean the cost of transport has gone up 10%, because you are also paying for the driver, maintenance on the truck, loading and unloading time, and (temporary) warehousing. Typically, fuel costs are 30% or less. Knowing this allows you to cap cost of service increases at 3% when fuel goes up 10%. Not knowing this will land you in hot water.

Furthermore, the best way to save money is to not spend it in the first place. Does the organization really need 500 new printers? If the organization has a green mandate and the only paper it needs to hold onto are signed contracts, then maybe it only needs 100 shared network printers in common areas. These printers might cost twice as much as regular desktop printers, but the organization would still save 60% by minimizing demand. Another great example is cleaning, as outlined in this paper. Chances are your office doesn’t need to be cleaned every night and your trash doesn’t need to hauled daily. Reducing demand 50% can reduce costs 50% with no negotiation whatsoever.
Be creative; think from both the supplier’s perspective and your own

Just because a supplier’s pricing is 10% above market does not mean that they will be able to come down by 10%. If they are a small, niche, supplier they may not be buying the volume necessary to get the discounts from their raw material providers that other, bigger, suppliers can negotiate. However, they may be able to offer deals on value added services that they can provide cost-effectively. For example, if your company does direct shipping to the end consumer, and they have invested in a state of the art packaging and labeling solution, maybe they can drop ship directly to your end customer through your 3PL (Third Party Logistics) and save you the 20% overhead of having the product go through a 3rd party distribution facility. Be creative and flexible. A good supplier will try to help you save money in a win-win situation.

Avoid disruptive initiatives if non-disruptive options are available

The last thing an organization should want to do is to replace a well-performing incumbent vendor even if a new vendor is 10%, 15%, or 20% cheaper. Switching costs are always higher than estimated and the ire you will generate among the various business units is rarely worth it. Use spend analysis intelligently and creatively to find ways to reduce incumbent costs instead, either through demand control, market-based restrictions on cost increases, or cost reductions on value added services that are cost effective for the supplier but cost prohibitive for you.

An analysis of the vendor invoices can point out errors in billing that might account for all of the variance in the incumbent’s pricing; these errors are not only correctable, but they result in immediate refunds or credits for the organization.

Demand flexibility from your spend analysis system

You should not have to wait 4 months, 4 weeks, or even 4 hours for a schema change. You should be able to make it almost instantaneously in your own copy of the data, to perform the analysis you want to perform when you want to perform it.

Rebuilding a cube of a few million transactions should not be an overnight task. In the worst case, it shouldn’t take longer than it takes you to get a cup of coffee. Literally. And you should be able to include data from your data sources of choice. AP. Your CM system. Your P2P system. Your contract labor management system. Whatever. Any data related to spend should be fair game.
Think short term (immediate ROI) or there may be no long term; but don't ignore the long term, either, because you'll get the funding to do it

Look for overpayments and categories where there is significant off-contract spend first. These will be quick hits. Detailed analysis that could result in switching suppliers or volume levels will take a lot of time and buy-in to accomplish. Settle for the 20% savings opportunities now, so that you will have the funding and time to go after the 80% savings opportunities later.

Have a plan, track progress, and continually communicate success

This cannot be understated. The best way to garner continued funding and support is to measure and show progress against a plan and communicate that regularly.

At the beginning of this paper, we indicated that gaining executive sponsorship is critical to success. If you don't help senior management to understand the progress you are making and the value that spend analysis and other sourcing tools is delivering, it's your fault. Make certain the key players know the facts of what you're done and how you've done it; and you'll have all the support you need for the next level of initiatives.
References

[Authors’ note: At the time of publication, active links to all of these online documents can be located on the first results page of a Google search for their titles]


[2] Spend Analysis Tools Help Companies Find Opportunities for Supply Chain Savings by Jean V. Murphy for SupplyChainBrain.com, October 2005


[5] Strategic Sourcing in the Mid-Market Benchmark by Tim A. Minahan, December 2005


[8] IBM Global Chief Supply Chain Officer Study: The Smarter Supply Chain of the Future by IBM, March 2009

[9] Success Strategies in Advanced Sourcing and Negotiations: Optimizing Total Cost and Total Value for the Next Wave of e-Sourcing Savings by Aberdeen, 2005


[12] PPG Analyzes, Then Brings Serious Focus to Managing Procurement by Jean V. Murphy, October, 2008


[15] The Intersection of Analytics and Supply Chain Risk Management - Using Intelligence to Drive Early Intervention
by Jason Busch, March 2010

[16] Using Sourcing Intelligence to Combat Commodity Volatility - The Price Index Advantage
by Jason Busch and Lisa Reisman, January 2009

[17] The Future of Procurement Analytics and Content - Three Scenarios for Company Adoption
by Jason Busch, February 2011

[18] Beyond Basic Scorecarding - Supplier Performance and Development Approaches to Drive Competitive Cost and Risk Advantages
by Jason Busch and Sherry Gordon, January 2011
Abbreviations

AI  Artificial Intelligence
AP  Accounts Payable
AR  Automated Reasoning
BI  Business Intelligence
ERP Enterprise Resource Planning
GL  General Ledger
GPO Group Purchasing Organization
MRP Manufacturing Resource Planning
MWBE Minority and Women-owned, Business Enterprises
P-Card Purchasing Card
PO  Purchase Order
SKU Stock Keeping Unit
UFC Ultimate Fighting Championship
UNSPSC United Nations Standard Products and Services Code