Chapter 10

Monitoring and Managing Data and Process Quality Using Data Mining: Business Process Management for the Purchasing and Accounts Payable Processes

Daniel E. O'Leary

Contents

10.1 Introduction

Recently, businesses have become more concerned about using transaction data to generate knowledge about the world in which they function, often referred to as so-called "business intelligence." This business intelligence typically is generated using tools such as data mining and knowledge discovery. Although much of that focus on business intelligence initially was generated about relationships with other firms, such as sales, increasingly there is a focus on internal processes. That focus of generating business intelligence about internal processes, to facilitate management and monitoring of those processes, is referred to as "business process management" (BPM).BPM can be used on any of a number of processes, such as sales analysis, accounts receivable analysis, inventory analysis, and other activities. However, this chapter focuses on purchasing and accounts payable processes so that particular metrics and approaches can be analyzed.

10.1.1 Purpose

BPM has received limited, if any, academic analysis to- ate but there has been substantial commercial development of BPM. Most commercial uses of BPM, particularly in accounts payable and purchasing, are aimed at a better understanding of payment activity and trends (Exhibit 10.1), and not at data quality or investigation of fraud. However, this chapter broadens that focus to examine data quality and consider how fraudulent activity might be spotted with BPM capabilities. Historically, when BPM and data quality are linked, it is a story that indicates how important data quality is to BPM. Unlike previous research, this chapter focuses more on how one can use BPM to monitor and ensure data quality.

Accordingly, the purpose of this chapter is to investigate how to assess and facilitate data quality, including spotting fraudulent activity. This is done using BPM as an organizing architecture for the data. The scope of this chapter is to investigate the processes of purchasing and accounts payable, within the context of business process management. As a result, this chapter focuses on the application of different approaches to facilitate data quality analysis of a process. Particular attention is given to data mining and its ability to ascertain when data seems appropriate or anomalous. There is also an effort to establish metrics that can be useful in facilitating the monitoring process.

10.1.2 This Chapter

This chapter proceeds as follows. While this section presents introduction to the chapter and its purpose, Section 10.2 summarizes some of 10.3 briefly reviews the specific domain of purchasing accounts payable and some generic sources of data quality disruption in those areas. Section 10.4 reviews the notion of BPM, while

Section 10.5 investigates metrics for purchasing and accounts payable that can be used to facilitate identification of data quality issues, such as fraud. Section 10.6 analyzes some approaches to determine how the data one sees matches up with what one would expect. Section 10.7 uses a data mining perspective to investigate the underlying data quality and how that data quality might be undermined by fraudulent data. Finally, Section 10.8 provides a brief summary of the chapter and its contributions.

10.2 Preventive and Detective Controls for Data Quality

The first step in ensuring data quality in virtually any setting is by using a strong set of preventive controls that prevent, to the extent possible, the entry of incorrect data or incomplete data. The second step is to build in additional controls that will facilitate detection of erroneous data or fraudulent data. In addition to characterizing controls as preventive or detective, controls also can be categorized as computer based or process based. This section investigates those control categories.

10.2.1 Preventive versus Detective Controls

Preventive controls are designed to limit errors or irregularities from being introduced into the data. A classic preventive control is a speed limit sign that indicates the upper bound on car speed. On the other hand, detective controls are designed to find errors or irregularities once they have been introduced to the data. A classic detective control is a radar gun that indicates how fast the car actually is going.

10.2.2 Computer-Based Controls

Computer-based controls use computer capabilities to provide control over the data quality. There are a number of computer-based controls that can facilitate data quality and control over a process, including the following.

10.2.2.1 Individual Accounts

Perhaps the most important control is the ability to have individual accounts for each user. This makes each individual directly responsible for the activity in their account. In these settings, each individual has his own password to control access over the account and corresponding purchases. Individual accounts allow for "virtual signatures," to indicate which user accessed the information and made the purchases, etc.

10.2.2.2 Drop-Down Menus

To ensure that the data entered comes from a feasible set, drop-down menus can be used to limit choice to a feasible set of entries. As a result, drop-down menus facilitate the prevention of bad data.

10.2.2.3 Forcing Completion of Specific Fields

To ensure that all of the necessary data is entered, the transaction can be held until all necessary data items are completed. Forcing completion provides a preventive control to ensure that all appropriate fields are completed and a detective control to find out when appropriate fields are not completed.

10.2.2.4 Forcing a Particular Type of Data

To ensure data quality, some fields may require a particular type of data. For example, some fields may require numeric data, while other fields may require alphabetic data.

10.2.3 Process-Based Controls

Process-based controls also can facilitate data quality and control. Rather than using technology capabilities, instead process control is attained by taking a few key process steps, building control into the process using the process or activities within the process.

10.2.3.1 Responsibility

An important process-based control is to assign responsibility for individual and process-based activities. If responsibility is assigned, then that person can provide a control, either detective or preventive, to make sure the data is correct. If there is a problem, responsibility can indicate who to tracked down to resolve the problem.

10.2.3.2 Separation of Responsibilities

To minimize the potential for fraud and error, key responsibilities can be separated. For example, in the discussion later, the responsibilities of the purchasing agent and the accounts payable clerk are "separated." The purchasing agent decides from whom goods should be purchased and generates the purchase order, while the accounts payable clerk is responsible for matching all the appropriate documentation needed to generate payment to the vendor. A third person is responsible for actually signing the check for vendor payment. By separating these responsibilities, there is increased control, and inappropriate behavior can be limited, unless there is collusion among the employees. Further, because we separate responsibility, individuals can detect problems by seeing others work. As a result, there is increased control and inappropriate behavior can be limited, unless there is collusion among employees.

10.2.3.3 Authorization

Authorization is a control that requires some individual to take responsibility for allowing a particular event. For example, large purchases typically require authorization by some appropriate level of management, whether it is a manager, the CFO, the CEO, or the Board of Directors, typically through a review and signature, either actual or digital. Authorization can prevent some errors because review allows one person the ability to detect errors, while the fact that someone needs to authorize an activity can serve as a deterrent to prevent unauthorized activity.

10.3 Purchasing and Accounts Payable

Purchasing and accounts payable require quality data because much of an enterprise's performance is based on the goods that it purchases. There are at least three scenarios that provide the basis to generate detailed key performance indicators (KPIs) and approaches. The analysis presented here spans these three different approaches.

10.3.1 Scenario 1: Classic Purchasing and Accounts Payable

In this first scenario, information flows primarily using documents. Purchasing processes typically are initiated internally by a "requisition," where a need for a purchase is established. Requisitions also provide preventive controls because a manager generally must have responsibility for authorizing the purchase. After the requisition is received, a purchasing agent establishes a purchase order that typically lays out the contract with a particular vendor to purchase the goods. Purchasing agents ensure that the vendors chosen are legitimate vendors and that the products that they provide meet certain standards. Purchasing agreements are sent to the vendor, receiving (so they know what to expect), accounts payable (responsible for payment), and purchasing agreements are kept in purchasing for reference.

Generally, after the goods have been sent by the vendor, an invoice is issued by the vendor and sent to the purchasing firm. When the goods are received, people in the organization's receiving department create a receiving memorandum. Typically, accounts payable gets a copy of the purchase order, the invoice, and the receiving memorandum; matches them; and pays the bill. This process is summarized in Figure 10.1.

Information is periodically digitized as documents are processed if there is a computer-based system supporting the process. For example, purchase requisitions

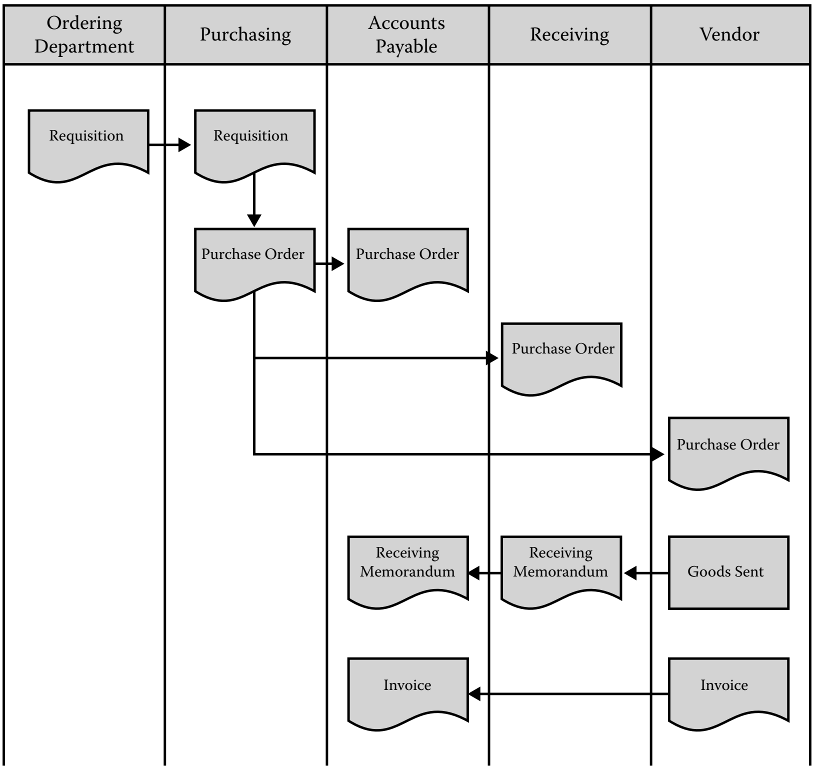

Figure 10.1 Classic accounts payable and purchasing.

could be created in digital format. Selected information from those forms could then be used to create a purchase order. It is likely that at least information about the vendor and the purchase are entered into the system so that, ultim(e.g., using electronic data interchange). This would facilitate single rather than multiple entries of the same data. In this latter setting, data would need to be input a single time for each document. At the other extreme, data in each functional silo must be input digitally with each document. In the second case, as an example, purchase order information would be entered into four different systems (purchasing, accounts payable, receiving, and at the vendor).

There are a number of sources of data quality disruption in this classic setting, including potential errors and fraud. Errors can originate from many sources. For example, data entry can generate input errors (e.g., data entry errors). The more times a document is entered into a system, the higher the likelihood of error. Purchasing can generate purchase orders with errors, and vendors can generate invoices with errors. Further, controls may not work or may be overridden. Thus, this chapter later reviews approaches designed to detect anomalous data to facilitate quality data.

10.3.2 Scenario 2: E-Purchasing

With the advent of E-purchasing, a number of companies developed intranet-based systems designed to facilitate and control purchasing. In these so-called "E-purchasing systems," purchasing typically arranges with different suppliers to provide digital catalogs from which system users can make purchases. Users are typically provided different "roles" (preventive controls) that indicate what kinds of goods they are authorized to purchase, (.g., office supplies, computers, etc.). In addition, users may have individual budgets for their total and individual purchases. The system then limits the kinds of purchases that they can make and controls the expenses that they can incur. The system also guides them to a preselected set of products that meet organizational constraints.

This approach provides firms with ongoing and summary digital information about purchases and purchasers. Ultimately, managers have responsibility to review the data and authorize the purchases. Unfortunately, in some settings, roles and budgets are not fully implemented or monitored and authorized. In those settings, individuals may exceed their purchases and may make purchases that are later converted to cash for their personal use. Accordingly, the lack of preventive controls may suggest that detective controls be used to supplement the control environment.

10.3.3 Scenario 3: Emerging Purchasing and Accounts Payable

In another scenario, involving decentralized organizations (e.g., universities), organizations are adopting or have adopted processes and technology that require less direct involvement by purchasing specialists, as purchasing activities are transferred directly to employees. In this setting, many of the process controls are sacrificed because of cost-benefit relationships. For example, for low-cost items, the actual purchasers may be in a position of selecting vendors without an extensive selection process. Receipts are used to get reimbursement because it is cost beneficial. However, in those situations, that may mean that the corresponding vendors are not legitimate vendors, that corresponding products may not meet quality requirements, or that the resulting transactions may be partially or completely fictitious. Again, detective controls can be used to supplement the control environment.

10.4 Business Process Management: Monitoring Data Flows for Purchasing and Accounts Payable Data Quality

Monitoring data and data quality increasingly has fallen under the auspices of so-called "business process management." Business process management (BPM) has a number of definitions; however, we use the following:

> BPM is the use of an integrated set of key performance indicators that are used to monitor an organizational process in real time. Business process management (BPM) is a management discipline that combines a process-centric and cross-functional approach to improving how organizations achieve their business goals. A BPM solution provides the tools that help make these processes explicit, as well as the functionality to help business managers control and change both manual and automated workflows.
>
> —**Microsoft** [7]

Effectively, BPM uses "business intelligence" approaches as a means of monitoring data streams. Data is obtained in real-time from sources such as an enterprise resource planning (ERP) system. Key performance indicator (KPI) metrics are used to summarize the data. Those metrics are then analyzed and some of them are presented to the appropriate managers for review, often in the form of so-called dashboards, to facilitate monitoring. If the data is anomalous, then the manager can act on the data in real-time. Further, increasingly, KPIs are being forecast to see if they are likely to become anomalous in the future.

10.4.1 BPM Dashboards

BPM dashboards use real-time data feeds to provide users with easy-to-use and easy-to-read measurement devices. Typically, classic green, yellow, and red colors are used to cast a corresponding interpretation of under control, borderline, and out of control, respectively. In the following example, it is easy to see that it is an example of each setting, without any specific numeric values. In this way, status quo is maintained and problems are quickly isolated. As an example, a three-dial dashboard is presented in Figure 10.2. Each of the dials is monitoring a KPI. The KPIs that they represent are, left to right, in control (green), on the edge of being in control (yellow), and out of control (red). Typically, color-based dials are used to ease interpretation.

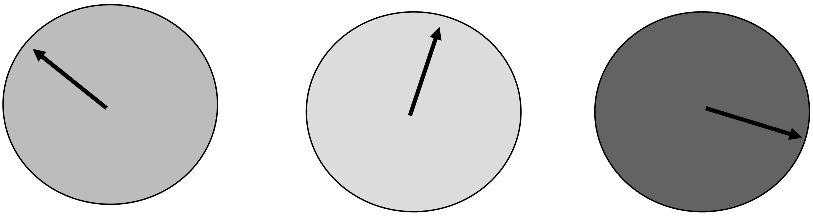

Figure 10.2 Three-dial dashboard system.

10.4.2 BPM Data Flows

In some cases, BPM data flows come from a single data source, such as an ERP system; however, in other cases they will come from disparate sources. In those settings, a consistent semantic mapping of the data will be necessary to ensure that the data and corresponding metrics are comparable. In some settings, this will prove one of the most important steps, and the BPM system will bring together disparate data flows under one system, sometimes for the first time. In some situations, a classic XML (eXtensible Mark-up Language) approach can be used to gather and label the data.

10.4.3 BPM Process Changes

However, in some situations, BPM is more than just capturing and managing KPIs on a particular process. In some cases, companies have changed the way they process invoices to facilitate BPM, particularly to meet the need for real-time data. Hasbro apparently developed a portal through which vendors could directly submit invoices [1]. After submission through the portal, invoices were routed to the appropriate vendor management teams for approval and from there for further processing. This approach increases the visibility of the approval and processing of the invoices, allowing them to better control the flow and understand bottlenecks, from time of submission to final payment.

10.4.4 Forecasts of KPIs

BPM systems may go beyond monitoring KPIs to actually monitoring forecasts of KPIs. Using real-time data, forecasts would be made and communicated to managers in a similar way as for real-time data. Forecast information would then be categorized as "in control," etc.

10.4.5 BPM Capabilities

What are key BPM capabilities? Historically, BPM takes data streams and puts them in a readable and accessible form so that managers can see critical data. As seen in Table 10.1, the focus has been on a range of important corporate issues. However, recently, BPM has been viewed as a potential tool for the analysis of fraud. For example, apparently the Louisiana Department of Social Services is now using BPM to facilitate identification of fraud [6]. This is one of the early applications aimed at using BPM to focus on issues other than productivity.

10.5 BPM Metrics: Purchasing and Accounts Payable

BPM metrics and systems can be used for different purposes. For example, BPM for accounts payable and purchasing can allow insight into cash outflows and facilitate cash planning. Historically, this means information about accounts payable

Table 10.1 Cognos Performance Applications Accounts Payable

Understanding Accounts Payable as Part of Financial and Supply Chain Analytics.

The pre-built reports and metrics of the Cognos Accounts Payable Analysis application give you a better understanding of your payment-related activity and trends. You can:

- Increase managerial productivity by reducing reporting and analysis time.
- See how much is due and when and the value of overdue accounts.
- Increase working capital by optimizing cash outflow strategies.
- Keep better control over cash outflow while maintaining strong vendor relationships.

Cognos Accounts Payable Analysis gives you more than 60 key performance indicators and more than 30 reports. These metrics and reports are grouped in four key areas of analysis, answering a variety of business questions:

- Accounts Payable Performance
  —What money is owed this period? What percentage is past due?
  —How quickly is the organization paying?
  —What percentage of accounts is not meeting terms? What is the value of overdue accounts?
- Accounts Payable Vendor Account
  —What is the current balance for a vendor account?
  —Which vendors are problematic? Why?
  —What is the cost to pay vendors, including errors, method of payment, and adjustments?
- Accounts Payable Cash Outflow
  —What is the expected cash outflow if no/all accounts take advantage of discounts?
  —What is the expected cash outflow based on the expected days to pay for each account based on payment patterns to date?
- Accounts Payable Organizational Effectiveness
  —How has account distribution across analysts changed as business has increased?
  —What was the total cost/savings for being in variance as related to payment terms?
  —What is the average/weighted average days past due?

Source: http://www.cognos.com/products/business_intelligence/applications/modules/payable.html that are outstanding, the accounts payable due to vendors, the extent of vendor discounts used, and the extent of overdue accounts. Table 10.1 provides summary of a BPM vendor's approach to purchasing and accounts payable, including goals and metrics.

However, historically, BPM has not focused much on ascertaining fraud and anomalous information. However, with the recent focus on the Sarbanes-Oxley Act, that focus could change. A number of metrics can be developed and monitored as part of a business process management system aimed at trying to find evidence of fraud or other data quality problems in purchasing and accounts payable. Some of those metrics include the following.

10.5.1 Number of Invoices Received from Suppliers

An important ongoing statistic is to capture the number of invoices received from each supplier on a monthly basis. Anomalous changes can indicate data quality problems. A steep increase or decrease in some vendor invoices may indicate that vendor numbers have erroneously been attributed to some invoices, for example, through data entry errors or a wrong vender number, whether purposefully or by accident. It may also indicate a fraudulent attempt by the vendor to obtain multiple payments.

10.5.2 Number of Transactions per System User

The system user varies based on the type of system in place, as discussed above. If one considers the number of transactions per accounts payable clerk, then the KPI provides a measure of productivity for the people involved in the accounts payable system. If one considers the number of transactions per worker using the system to buy goods, then the number of purchases can represent time spent away from their job, and also provide insight into how much productivity is spent on such issues.

Anomalous changes from month to month can indicate data input errors or fraud for at least two reasons. First, it may be that the wrong user is attributed to the transactions. An inappropriate user may be masquerading as another user. Second, if the user is replaced, then the replacement is likely a new user, and higher error rates are attributed to new users.

10.5.3 Percentage of Invoices Paid without a Purchase Order Reference

A classic fraud approach is to send goods and then invoice for them although the goods have not been ordered. This approach typically charges a premium price for substandard goods. As a result, firms often require a purchase order, as seen above in our three scenarios.

Further, although a preventive control is to require a valid purchase order number, in some systems without the proper controls there may no purchase order required to be associated with an invoice. The lack of a purchase order can indicate that the transaction is fraudulent or in error, but in any case anomalous.

10.5.4 Number of Invoices for a Purchase Order

Knowing that a purchase order number may be required, users in the process of doing a fraudulent transaction might use a legitimate purchase order number as part of the data input process, but one that is not appropriate for the particular invoice. In that situation, there are likely to be multiple invoices for a purchase order number. Thus, a list of the higher numbers of invoices per purchase order could be a KPI of interest.

10.5.5 Number of Users Using Each Vendor

In some cases, the number of users of a vendor can be indicative of a data quality problem such as fraud. For example, if a user and a vendor are working to defraud the company, it may be that the user would be the only one in the firm affiliated with that vendor.

10.5.6 Relative Size of an Invoice

There are a number of stories of organizations putting the decimal point in the wrong place on a payment, so that a $ 100 payment becomes a $10,000 or larger payment. Accordingly, a major concern is that the dollar amount of a payment, not be excessive. There are a number of tests for ascertaining anomalies. One such test is the ratio of the largest payment to the second-largest payment (e.g., [8]). This can be generalized to the ratio of the j-*th* largest payment to the (j + 1)*st* largest payment. Whenever that ratio is substantial, it can indicate a problem with data quality and may be indicative of fraud. In the case where fraud was being purposefully committed and there was awareness of the existence of a test comparison between the first and second payment sizes, two large fraudulent transactions could be executed, thus mitigating the effectiveness of that test. As a result, comparison of more than the first two adjacent invoices would be appropriate.

10.6 Knowledge Discovery: Comparison to Expectations

Although preventive controls are critical to ensuring that data quality is high, an important approach to ensuring data quality is to compare data to "expectations" to see if the data meets those expectations. There are a number of bases of comparison, including Benford's law and other comparisons.

10.6.1 Benford's Law

One metric that can be traced and monitored to expectations is Benford's law [4, 10], which states that the first significant digit d ($d \in \{1, ..., b - 1\}$) in base b ($b \geq 2$) occurs with probability proportional to $\log_b(\mathrm{d} + 1) - \log_b(\mathrm{d})$.

As a result, Benford's law establishes a set of expectations for the distribution of numbers. For many numeric generating processes, the first digit (or first and second, etc.) can be analyzed to see if it meets expectations. If it does not, then that can indicate an anomaly and that an investigation should be conducted to determine if there is some fundamental problem.

Numeric sequences could include a wide range of information; for example in State of Arizona v. Wayne James Nelson (1993) [8], the accused was found guilty of attempting to defraud the state of roughly $2 million. A manager in the Arizona State Office of the Treasurer had diverted funds to a bogus vendor. The amounts of the 23 checks issued are shown in Table 10.2. The first digits are almost all 8 and 9,

Table 10.2 Data from *State of Arizona* **v.** *Wayne James Nelson* **(1993)**

Date of Check	*Amount($)*
October 9, 1992	1,927.48
	27,902.31
October 14, 1992	86,241.90
	72,117.46
	81,321.75
	97,473.96
October 19, 1992	93,249.11
	89,656.17
	87,776.89
	92,105.83
	79,949.16
	87,602.93
October 19, 1992	96,897.27
1	91,806.47
1	84,991.67
1	90,831.83
1	93,766.67
1	88,336.72
1	94,639.49
1	83,709.28
1	96,412.21
1	88,432.86
1	71,552.16

the numbers that one would expect to see the least of. As a result, compared to Benford's law, the results are anomalous.

Not only can Benford's law be used to see if there is potential fraud, but it might also suggest that policies are not being followed or that particular policies are being avoided. For example, if there is a policy that expenses must be signed when they exceed a certain amount, an analysis of the data is likely to find an abnormal number of expenses filed just below the threshold. For example, if there is a cut-off of $500.00, where expenditures at $500.00 must be signed, there is likely an abnormally large number of expenditures around $400.00.

10.6.2 Accounts Payable Data Analysis

Given a database of accounts payable data, a number of comparisons between the data can be made to help establish the quality of the data. Three key data elements in accounts payable and purchasing are the invoice number, the amount of the invoice, and the vendor number.

10.6.2.1 "Same, Same, Same"

An important test of the quality of the accounts payable data is for duplicate payment of the same invoice to the same vendor (see, for example, [8]). In this situation, the data is investigated for the same invoice number, same amount, and same vendor. Such duplicate payments can occur if the vendor provides multiple copies at different times of the same invoice, whether as part of normal business practice or as part of a fraudulent approach. As part of the analysis, the accounts payable clerk ultimately responsible for the match must be determined so that it can be ascertained if there is a systematic problem.

10.6.2.2 "Same, Same, Different"

One test of the quality of accounts payable data is the "same, same, different" test (same invoice number, same amount, different vendor) (see, for example, [8]). The purpose of the test is to compare different accounts payable entries to determine if they are the same, and as a result, a bill has been paid more than once or if the wrong vendor has been paid. An invoice might be paid twice in the situation where the invoice was paid to the wrong vendor and then the correct vendor. The wrong vendor may have been paid, either purposely or by accident, such as an incorrect keying of the data. As part of the analysis, the accounts payable clerk ultimately responsible for the match must be determined so that it can be ascertained if there is a systematic problem.

10.6.2.3 "Same, Different, Different"

Another test of accounts payable data is the reuse of a purchase order number for other amounts or vendors. In a system that requires a purchase order number, a fraudulent entry could "reuse" a purchase order number to meet the need of providing a purchase order number with each entry. This test would allow detection of such reuse.

10.7 Data Quality-Based Data Mining

Purchasing and accounts payable systems depend on the underlying data in the system being "good data" to begin with. However, that assumption may not be true. One approach to analyzing data quality is to investigate the data using data mining, in order to determine if the basic data set contains any anomalies, indicating problems with the underlying data. For example, vendors may be fraudulent or goods may be bogus, in which case any transactions involving those vendors or goods would be suspect.

10.7.1 Determining "Inappropriate" Vendors

After the attacks on the World Trade Center, in New York City on September 11, airlines in the United States began comparing airline passenger lists to so-called "bad guy lists" for use in systems such as "NORA" (Non-Obvious Relationship Awareness). These systems were designed to find passengers who might be terrorists.

If the vendors that an organization does business with are not appropriate, then the data generated in interaction with those firms may be lacking quality, and the transactions may be fraudulent. As a result, similar to NORA, firms could compare their own employee, vendor, and customer lists to "bad guy lists" (BGLs) to facilitate determination as the "appropriateness" of employees, vendors, or suppliers. In some cases, a broader-based approach might be taken by including in that comparison incident and arrest systems. These comparisons are detective controls that may determine inappropriate vendors that were not prevented from being part of the system at the beginning.

10.7.2 Determining Fraudulent Vendors

Generally, vendors are third parties that operate at "arms length" from the particular organization. Data mining can be used to determine if there are any fraudulently created vendors. Vendor characteristics can be matched to characteristics of other agents associated with the organization. For example, vendors' characteristics can be compared against employees, because it would be rare that an employee would also be a vendor. If employees were vendors, then it would at least be of enough concern to enumerate and examine. Agent characteristics such as name, address, phone number, or even bank account numbers could be compared for similarity in the different databases.

10.7.3 Fraudulent Company Shipment Addresses

Products are "shipped to" particular addresses as part of purchase agreements. Generally, those "ship to" addresses are from a subset of organization addresses where the particular organization does business. As a result, if a "ship to" address does not come from that set, then it might indicate a fraudulent transaction and definitely would be anomalous. This likely would be even more indicative of a problem if the "ship to" address corresponds to an employee address. This is not to say that all such shipments would be suspect; for example, there may be a home office. However, such a correspondence between addresses could indicate fraudulently obtained goods.

10.7.4 Selected Issues in Comparison of Vendors

The analysis of shipment and vendor data could be done by people, but generally, using an intelligent system would be faster and possibly more effective, given the nature of the task. Such comparisons could take some intelligence to execute well. First, name information may be inconsistent. For example, "International Business Machines" may be in the database under that name or "IBM" or "I.B.M." or any of a number of other alternatives. Second, address conventions may be inconsistent. For example, at some point in addresses, "N." would need to be considered the same as "North" and "E." would need to be considered the same as "East." Similarly, other abbreviations, such as "St." would need to be processed as "Street." Third, phone number information may be non-standard. For example, phone numbers could include dashes or not include that information. All of these issues would limit the ability of a system to correctly match vendors in different systems.

10.7.5 Bogus Goods

Left uncontrolled, users could conceivably order goods, have their company pay for them, and then resell the goods. For example, computer memory chips can be ordered by individuals in Scenario 3, likely from a vendor of their choice. It would be possible to contrive such purchases where the user was able to transfer the money from the company to himself, as long as each individual purchase and the purchases in aggregate did not exceed some amount. One approach to detect this kind of behavior is to keep track of different kinds of goods and how many of each kind each user orders.

To better control such purchases, additional preventive control information about goods could be specified. Goods could be characterized as "limited consumption goods." Whenever goods purchased exceed a certain amount, the purchases could kick out as anomalous. For example, computer memory chips could be characterized as a limited consumption good, where purchases for that type of good should not exceed some particular limit.

10.8 Summary and Contribution

This chapter investigated approaches to ensure and analyze data quality in a purchasing and accounts payable process, in the context of business process management. It summarized different types of controls, preventive and detective, and their use in computer systems and processes. Further, three different scenarios of how the purchasing and accounts payable processes would be generated were analyzed, as the particular implementation indicates what limitations are likely. Then the notion of business process management (BPM) was introduced. BPM provides a renaissance of managing processes, by integrating technology into that management process.

The primary contribution of this chapter is the development of an architecture for the use of business process management to analyze data within purchasing and accounts payable for data quality and potential fraud. Historically, BPM has not been aimed at those activities but has focused more on managing cash flows in the process. This was done by laying out some metrics to monitor process data, discussing how knowledge discovery could be used to determine if data is meeting expectations, and how data mining could be used to investigate the data quality of the underlying system information.

References

1. Chen, A. "Hasbro Plays to Win with BPM," eWeek.com, August 2, 2004.
2. Cognos, "Cognos Financial Analytics," http://www.cognos.com/pdfs/whitepapers/wp_cognos_financial_analytics.pdf
3. Coderre, D. "Global Technology Audit Guide Continuous Auditing," The Institute of Internal Auditors, 2005.
4. Hill, T. "The first digit phenomenon," *American Scientist* 86 (July–August 1998), p. 358. http://www.americanscientist.org/template/ AssetDetail/assetid/ 15660; jsessionid= baa6gWCz81?fulltext=true
5. Lombardi Software, "Accounts Payable," http://www.lombardisoftware.com/bpm-accounts-payable.php#
6. Metastorm, "The Louisiana Department of Social Services," 2006, www.metastorm.com/customers/lodss/Louisiana%20DSS%20Success%20Story.pdf

7. Microsoft, "Business Process Management Overview," http://www.microsoft.com/biztalk/solutions/bpm/overview.mspx
8. Nigrini, M. "I've Got Your Number," *Journal of Accountancy*, May 1999. http://www.aicpa.org/pubs/jofa/may1999/nigrini.htm
9. Potla, L. "Detecting Accounts Payable Abuse through Continuous Auditing," *IT Audit*, The IIA, Altamonte, Springs, FL, Vol. 6, November 2003.
10. Wikipedia, Benford's Law, http://en.wikipedia.org/wiki/Benford's_law