KPMG FORENSIC

Six Sigma in the Legal Department

Obtaining Measurable Quality Improvements in Discovery Management

ADVISORY

AUDIT ▪ TAX ▪ ADVISORY

"If you can't measure it, you can't manage it."

— Axiom attributed to Peter Drucker and others

Six Sigma is a methodology intended to improve performance in defined processes in relationship to an aspirational goal. Responsibility for the lawful, efficient, and effective response to subpoenas and other legal process in a particular legal proceeding, including decisions regarding how to respond to such process and what to produce and what not to produce in response to such legal requests, is the responsibility of counsel through their exercise of professional judgment applied to particular legal issues relevant to a specific legal proceeding, and with due consideration of controlling law, court rules, and discovery orders in effect in the specific legal proceeding. The Six Sigma processes discussed in this white paper constitute a methodology designed to measure the efficiency and cost-effectiveness of document preservation/collection, paring, processing, review, and production by counsel and are not intended to establish or advocate an independent standard by which to judge a party's compliance with any legal obligation with respect to document collection, review, or production. Application of these processes to document preservation/collection, paring, processing, review, and production will vary from proceeding to proceeding, depending on the unique circumstances of each proceeding. The methodologies expressed herein do not constitute legal advice or the practice of law; and application of Six Sigma methodologies to a litigation process is not a guarantee of any specific level of accuracy, cost-effectiveness, or financial savings. Nothing herein should be understood to be advising any party to litigation or other legal proceeding as to how the party should satisfy its legal obligations in the context of a specific proceeding or in response to a specific subpoena or document request.

Foreword

Over the past five years, law departments and law firms have seen a substantial rise in the volume of electronic discovery. A 2005 study released by Fulbright & Jaworski LLP found that "electronic discovery was the number one new litigation-related issue for companies with revenue over $100 million." Further, it reported that, "the issue general counsel wants the outside counsel to understand the most is cost control."[1] According to a survey of chief executive officers (CEOs) released in the October 2005 issue of *Corporate Legal Times*, the most important thing general counsel (GC) could do to improve the legal department is "communicate more with business units."[2] It appears that law departments increasingly should expect to be measured by the same business metrics as the rest of the organization.

Industries such as energy, financial services, manufacturing, pharmaceuticals, and technology have responded to similar levels of cost concern by adopting widely recognized quality improvement and customer service standards. Six Sigma®, a set of standards developed by Motorola and embraced most famously by GE, has become the preeminent quality improvement methodology today.

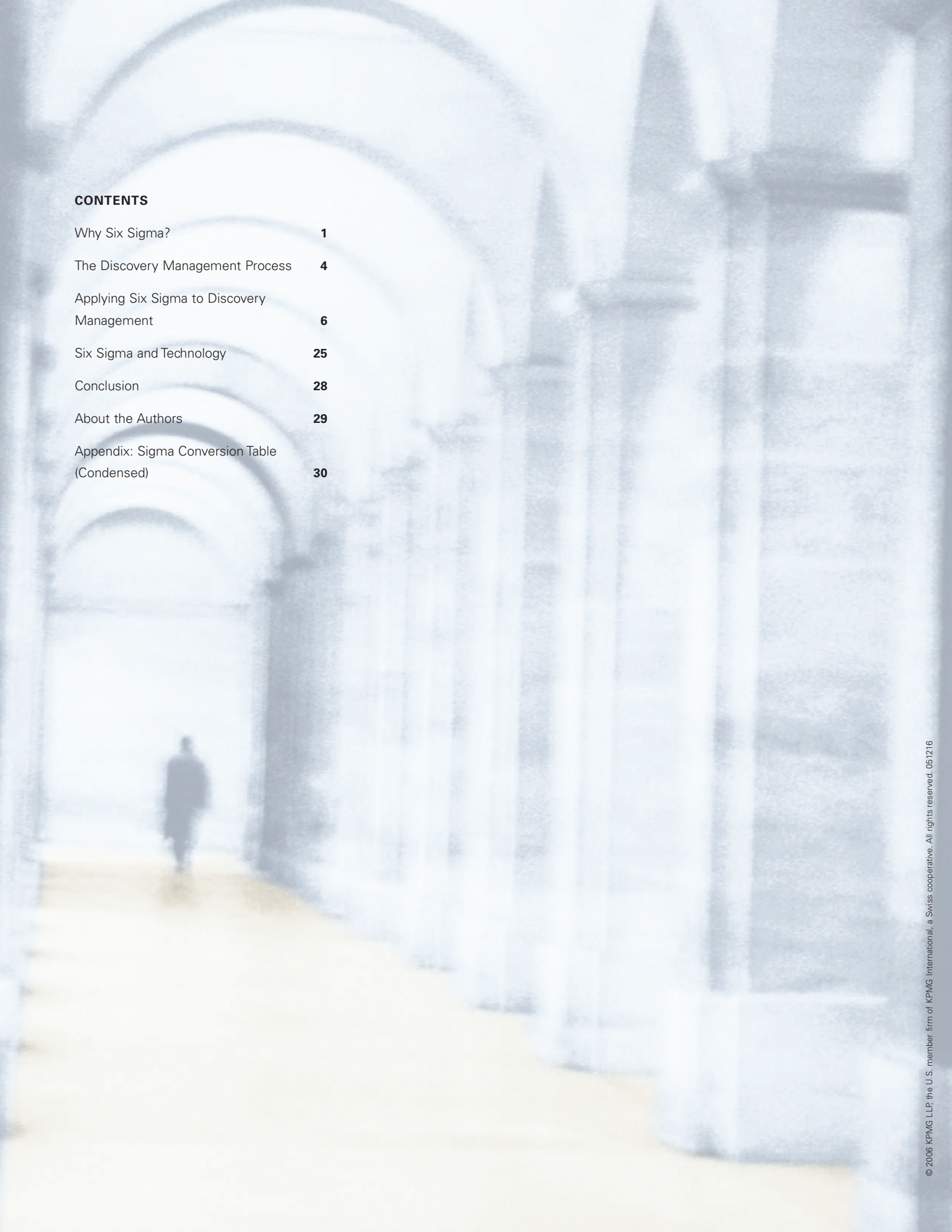
Following upon an earlier KPMG white paper—*A Revolution in e-Discovery: The Persuasive Economics of the Document Analytic Approach*[3]—which analyzed e-discovery costs, this paper, a joint effort by KPMG Forensic℠ and Pivotal Resources, Inc., explores how law firms and law departments can apply Six Sigma process improvement methodologies and tools to help meet increasing demands for cost control and efficiency in the discovery process for litigation and investigations. This paper focuses on e-discovery because it is usually the largest cost center of litigation and investigations.

[1] "Second Annual Litigation Trends Survey," 2005, Fulbright & Jaworski LLP

[2] *Corporate Legal Times* / Dickstein Shapiro Morin & Oshinsky Survey of CEOs, "The Bottom Line," Cathleen Flahardy, *Corporate Legal Times*, October 2005

[3] KPMG LLP, 2005

CONTENTS



Why Six Sigma?

Six Sigma has enabled some of the world's largest corporations to reduce costs and increase efficiency while improving customer service. Six Sigma companies include Bristol Myers-Squibb, Citicorp, Ford Motor Company, GE, Honeywell, Lockheed Martin, Motorola, Sears, Sony, Starwood Hotels, and many others.

While there is no single definition of Six Sigma, or a single approach to implementing it, the following provides a useful summary. Six Sigma includes:

- A *management system* to achieve lasting business leadership and superior performance
- A *goal* for improvement that strives for near perfection
- A *statistical measure* to define the capability of a process

Six Sigma as a Management System

Six Sigma is not a single tool, but a business philosophy that focuses on a disciplined, data-driven, and measurable way of operating and applies many kinds of statistical, business, and operational tools in a meaningful, logical way.

The Six Sigma philosophy has had broad application across industries and has equal relevance to the legal profession. Corporate law departments—often viewed as a necessary cost of doing business—can benefit from measuring such important functions as "process management," "efficiency," "process improvement," and "cost savings." These issues resonate in the minds of GCs, CEOs, chief financial officers (CFOs), and other officers of Six Sigma–influenced organizations. Outside counsel can benefit from understanding and adopting the Six Sigma approach to help their lawyers speak the same language as their clients from a problem-solving perspective, provide higher-quality deliverables, and improve customer service.

Six Sigma as a Goal

In any business, a logical goal is to minimize the number of defects over time or units in a given process. In professional service organizations, a "defect" occurs when key customer requirements are not met. It is generally accepted that defects in a manufacturing process result in higher costs and lower profits. Defects in service and transactional processes—that is, unmet customer expectations—have the same effect, although sometimes they are not as obvious.

With Six Sigma, the goal is to redesign or improve the process so that defects do not occur in the first place. For example, Six Sigma teams focus not on "building a better mousetrap," but rather on why the mouse is appearing in the first place. Companies operating under Six Sigma strive for products and services that are essentially defect-free.

As a reference, the following abbreviated table focuses on DPMOs (defects per million opportunities), a measure of how many mistakes occur over one million activities. The table also identifies the "yield" percentage, which reflects the proportion of work done correctly, that is, with no defects. The chart lists the range of sigma (σ) rankings with corresponding DPMOs and the percentage of work done right (yield).

Yield	DPMO	Sigma (σ)
99.9997%	3.4	6
99.977%	233	5
99.379%	6,210	4
93.32%	66,807	3
69.2%	308,537	2
31%	690,000	1

See Appendix for a more detailed version.

Given the data in the table, consider some real-world examples in a legal department or law firm setting:

- For every 300,000 documents reviewed for privilege…
  - *With 99 percent yield (approximately 3.75 σ):* ***3,000 missed documents***
  - *With 99.9997 percent yield (6.0 σ):* ***1 missed document***
- For every 500,000 computer-based review hours…
  - *With 99 percent yield (approximately 3.75 σ):* ***5,000 downtime hours***
  - *With 99.9997 percent yield (6.0 σ):* ***fewer than 2 hours of downtime***

In all cases, the practical difference between 99 percent quality and Six Sigma is staggering. Yet the downside of *not* pursuing ambitious quality goals can be equally staggering, as demonstrated in well-known cases[4] where defendants were either sanctioned or received an adverse inference jury instruction for mishandling electronic evidence.

Six Sigma as a Statistical Measure

Before 1987, when Motorola began developing its methodologies, sigma—the lowercase Greek letter σ—had little meaning outside the field of statistics. In statistics, σ is a symbol for "standard deviation," defined as the amount of variation within a set of data, a set of items, or a process.

Take, for example, a production review team of 40 attorneys, comprising 20 in-house attorneys intimately familiar with the case and 20 outside (contract) attorneys with little case familiarity. Each attorney is charged with reviewing a set of 200 documents at a time.

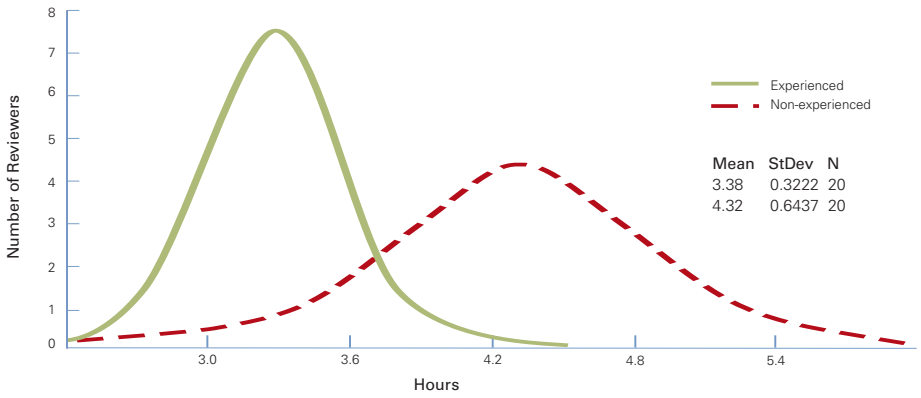
Comparing the two groups, the contract attorneys take longer on average to complete the production assignment since their lack of familiarity with the case details usually means they make document decisions more slowly. In addition, the *range* of average times spent

[4] *See United States v. Philip Morris USA Inc.,* Civil Action No. 99-2496 (GK); *Coleman (Parent) Holdings, Inc. v. Morgan Stanley & Co. Inc.,* CA 03-5045 AI; *Zubulake v. UBS Warburg LLC,* No. 02 Civ.1243 (SAS)

by the reviewers—the standard deviation—is more dispersed with the contract attorneys as compared with the in-house attorneys. Contract attorneys usually lack a collective understanding of the details of the case, which forces them to review *at their own pace and comprehension level*—some fast, some slow. On the other hand, in-house attorneys make document decisions not only more quickly but also at rates that are more consistent with other in-house attorneys. Having been trained in the specifics of the case, they share an understanding of the issues and have less cause for performance variations—standard deviation—within their group than do contract attorneys.

We can demonstrate standard deviation visually (see chart below) by plotting the average review times for the in-house attorneys (solid line) and contract attorneys (dotted line) by reviewer. The legend shows that the case-experienced, in-house group averages 3.38 hours per review set, with a standard deviation of .3222, while the contract attorney group averages 4.32 hours, with a standard deviation of .6437.

TIME TO REVIEW A PRODUCTION SET OF 200 DOCUMENTS *(Normal)*



Because Six Sigma focuses on standard deviation of performance rates rather than on absolute performance rates, it is not restricted by units of measure or business types. This approach gives Six Sigma universal applicability across industries. In addition, by basing its metric on what is important to the customer, it sets a higher standard than that set by non–Six Sigma companies.

In Six Sigma, "A 90 percent to 98 percent success rate does not deserve an 'A' rating. Rather, Six Sigma methodologies strive for a 99.9997 percent success rate, or a mere 3.4 defects per one million opportunities."[5] Activities that can be measured precisely—such as TIFF imaging—are highly conducive to Six Sigma analysis and improvement. On the other hand, activities in which some errors may be highly subjective—such as litigation—may still lend themselves to the Six Sigma methodology, but may be resistant to the highest sigma ratings. Nevertheless, as long as companies are "increasing their sigma" over time they are operating under the Six Sigma "doctrine."

[5] *The Six Sigma Way: How GE, Motorola, and Other Top Companies Are Honing Their Performance*, by Peter S. Pande, Robert P. Neuman, and Roland R. Cavanagh. Copyright 2000 by The McGraw-Hill Companies, Inc.

The Discovery Management Process

To understand how Six Sigma applies to discovery management, we need to understand the phases of discovery. We have organized the process in five phases—the 5 P's.[6] Your process owner may have different designations for discovery phases, but the message is the same: to apply Six Sigma to discovery you need to break up the process into discrete parts.

Plan: Scope Definition and Project Planning

Every discovery process should begin with proper planning and scope definition. During the plan phase, you need to identify the custodians, determine the locations of relevant information for each custodian (e.g., hard drives, network shares, PDAs), and estimate the magnitude of the data (in pages for paper documents and in gigabytes for electronic data). Then, using this information, you can identify the best tools and technologies available and assemble a budget and action plan.

Preserve: Digital Evidence Collection, Preservation, and Recovery

Once the project plan and budget are approved, a data-preservation team—consisting of company personnel, forensic technology personnel, or a combination thereof—executes a predefined collection plan. Special consideration should always be given to chain-of-custody[7] tracking and protection of evidence from spoliation.[8]

Pare: Catalog, Dedupe, Filter, and Analyze File Content

A gigabyte of e-mail can yield between 50,000 to 100,000 pages (75,000 pages for estimation purposes) and the typical custodian can generate between one and two gigabytes of data for review, even after some deduplication and keyword filtering. Consequently, a project involving 20 custodians, each with two gigabytes of data, could generate approximately three million pages, the equivalent of 1,500 banker's boxes of paper.

Old-fashioned methods of reviewing evidence for production are ineffective in dealing with this volume of data. Review teams must find ways to reduce the volume early in the process. "Document analytic" technologies help to reduce the volume by enabling professionals to use computers and algorithms to analyze the relevance of individual documents within various contexts and thereby significantly improve document review efficiency.

[6] *A Revolution in e-Discovery: The Persuasive Economics of the Document Analytic Approach*, KPMG LLP, 2005

[7] **Chain of custody** is a concept that applies to the handling of evidence and its integrity; it also refers to the document or paper trail showing the seizure, custody, control, transfer, analysis, and disposition of physical and electronic evidence.

[8] **Spoliation** refers to destruction of evidence relevant to a legal proceeding. The spoliation inference is a negative evidentiary inference that a finder of fact can draw from a party's destruction of a document or thing that is relevant to an ongoing or reasonably foreseeable civil or criminal proceeding.

By increasing "document decisions per hour," reviewers can quickly pare or reduce a voluminous data set while it is still in digital format. By the end of the paring phase, the potentially responsive data set is typically reduced by 80 percent or more. This is all done prior to any conversion to TIFF or PDF formats.

***Process:* Metadata and Text Extraction, Indexing, and Optional Conversion to TIFF/PDF Format**

Once the potentially relevant documents are identified and flagged, this subset is then processed (i.e., metadata and text are extracted and indexed) and often loaded into an online discovery management application.

***Produce:* Online Search and Review**

Once the loading is completed, the review team creates folders, delegates and administers access rights, and facilitates additional levels of review and quality control (e.g., verifying issue coding carried over from the paring phase). Further, selected documents can be "TIFFed" on demand for redaction, bates numbering, or other types of branding.

Applying Six Sigma to Discovery Management

Six Sigma uses two measurement-based methodologies to improve processes and reduce variation. One of these methodologies, DMAIC (pronounced duh-MAY-ick), focuses on improving existing processes, such as the 5 P's.[9]

Given that each litigation matter has its own set of challenges, deadlines, strategies, and technology considerations, discovery for each matter should be treated as its own DMAIC project. Under this methodology, the legal team needs to:

- *Define* the problem, the scope, and key issues, risks, deadlines, and customer needs for the discovery review and how the legal team plans to address them
- *Assess* what measurement tools will be used on the project and how data will be gathered
- *Analyze* the data to identify potential root causes and track "as is" performance
- *Develop* improvement plans to address key issues
- *Develop* a control plan for maintaining the improved results over time.

DEFINE

This first DMAIC step in the context of litigation discovery—defining key discovery issues—often poses the greatest challenge. Common topics include identifying and selecting the key issues to address; identifying the business case, constraints and assumptions, pain thresholds, the customer, and the customer's key requirements; documenting how work is currently being done; and assessing the benefits of making particular improvements.

From a discovery review perspective, the "customer" can be both internal and external, including:

- Senior company management
- Reviewing attorney(s), whether internal or external
- Opposing counsel or a government agency requesting documents.

Following are some Six Sigma tools and templates that contain hypothetical e-discovery issues that can assist your law department or law firm in the Define step.

[9] The second methodology, DMADV, focuses on developing new processes: define, measure, analyze, develop, and validate.

Project Evaluation

Project evaluation is an opportunity to anticipate resource problems and make adjustments based on needs. Accordingly, it is important to take a holistic view of the e-discovery process. You should assess your company's or your department's capabilities in all of the 5 P's: planning, preserving, paring, processing, and producing. For example, problems related to budgeting and scope may be caused by inadequate planning measures. Problems with chain of custody or spoliation of evidence may relate to preserving. Problems with inefficient, overwhelming, or expensive document review may be related to how your firm or company is handling the paring phase. Problems with metadata or file processing might relate to the processing phase, and problems with discovery review, issue coding, document production, redaction, and so on, would fall under production.

When identifying the pain thresholds, some helpful project selection criteria[10] include:

- Impact of business strategy
- Financial benefit, immediate and long term
- Urgency
- Risk management
- Resources needed
- Expertise available
- Organization or management buy-in
- Likelihood of success
- Learning or experience opportunity

DMAIC Charter

Once a project is selected (or accepted), the DMAIC charter is created (see the example charter on the next page). This charter serves as the blueprint or road map for your discovery task. The charter will vary according to your specific needs, but generally includes:

- *Business case*. Why is this opportunity or challenge being chosen?
- *Problem/goal statement*. What is the specific problem or pain that will need to be addressed in this litigation and what results will be sought from a discovery perspective?
- *Constraints/assumptions*. What time or resource limitations are placed on this project?
- *Scope*. How much of the process and range of issues is "in bounds"? For example, what items will internal versus external counsel manage?
- *Players and roles*. Who are the team members? Who is the sponsor or champion? Who are the stakeholders and who are the adversaries? Who is the customer and what does the customer need?

[10] *The Six Sigma Way: How GE, Motorola, and Other Top Companies Are Honing Their Performance*, by Peter S. Pande, Robert P. Neuman, and Roland R. Cavanagh. Copyright 2000 by The McGraw-Hill Companies, Inc.

EXAMPLE DMAIC PROJECT CHARTER

Business Case	To meet the challenges of our firm's XYZ litigation matter, we need to improve our processes and technologies around electronic data processing and document review. Given the nature of this government investigation and the tight deadlines imposed on the firm, increasing our review team capabilities and quality will be mandatory in order to meet our production deadline of four months.
Problem Statement	Our last electronic discovery engagement was challenged by too many error corrections and rework that resulted in a loss of $120,000 worth of billable time write-offs coupled with a negative client experience.
Goal Statement	Reduce reviewer defects by 80 percent by identifying and addressing root cause(s).
Constraints and Assumptions	Constraints: • Time constraints in double-checking review team's document decisions—XYZ matter has a four-month production timeline • Budget constraints in technology and people • Team members work no more than 25 percent of time on project • Solution should work on multiple engagements Assumptions: • Project leader/sponsor meets with team at least once a week. • Management will support and approve up to $200,000 in investment capital, as determined by DMAIC team.
Scope	Our project will analyze document review decisions on the XYZ matter for purposes of improving our discovery management quality control. Legal strategy and trial preparation will not be addressed in the context of this project.
Players/Roles	**Team Leader:** <Name> **Team Members:** Sponsor/Champion: <Name> Process Owners: <Name 1> <Name 2> **Resources:** List available resources or needed departmental cooperation (information technology, human resources, corporate, etc.).

Preliminary Plan

On completion of the charter, a preliminary plan (see sample below) should be written. This document addresses when each phase (D, M, A, I, and C) will be completed and by whom.

SAMPLE PRELIMINARY PLAN

ACTION/MILESTONE	WHO'S RESPONSIBLE	TARGET COMPLETION	√	ACTUAL COMPLETION
DEFINE				
Complete draft of charter	J. Smith & M. Jones	March 1	√	March 2
Complete voice of customer analysis	J. Smith & M. Jones	March 5	√	March 5
Complete process identification	J. Smith & M. Jones	March 20	√	March 20
MEASURE				
Prepare data collection plan	B. Doe	April 1	√	March 30
Complete initial data collection	B. Doe & M. To	April 30	√	April 30
Calculate current sigma	B. Doe & M. To	May 2	√	May 2
ANALYZE				
Confirm process improvement strategy	J. Smith	May 5		
Develop initial cause hypothesis	J. Smith & M. Jones	May 10		
Complete cause verification	J. Smith & M. Jones	May 15		
IMPROVE				
Develop draft solution list	Team	May 25		
Select solution for piloting	Team	May 30		
Pilot implementation complete	J. Smith & M. Jones	June 15		
Full implementation complete	J. Smith & M. Jones	June 30		
CONTROL				
Measure ongoing productions	B. Doe	July 30		
Documentation complete	J. Smith	July 30		

Voice of the Customer

In Six Sigma, determining who is the customer and what the customer needs is called "listening to the voice of the customer" (VOC). Translating VOC requirements can sometimes be a difficult task and is especially so in a litigation discovery context. As mentioned earlier, the customer in a discovery management context can include multiple parties, such as management, external counsel, and (albeit, counterintuitively) even your adversary in the litigation. The customer may be inside or outside the company, but is by definition *outside* the discovery management group.

With such a potentially diverse set of customers, it is particularly important to translate VOC into specific requirements. See the table on the next page for examples.

TRANSLATING VOC REQUIREMENTS

CUSTOMER SAYS	MEANING TO LEGAL TEAM	CUSTOMER REQUIREMENTS
We are spending too much on electronic discovery processing.	Management sees us as not effectively managing the electronic document review.	Management requires weekly updates and progress reports on the status of electronic document review that specifically measure document review efficiency and document decisions per hour.
The review process is going too slowly.	We are seen as not being efficient.	Customer requires a minimum of 1,500 document decisions per day per reviewer.
You don't have control of our documents.	Our discovery management system is not tracking documents in a way that meets our processes and chain-of-custody needs.	Each document record needs to contain a "source" file path field that shows the original source of the file.
Your document production is inconsistent across multiple jurisdictions.	The company is being sanctioned by the courts because it is failing to make consistent document decisions across multiple law firms.	Need to track document production history at the "document" level to ensure that previously withheld documents don't get produced in future productions.

Process Mapping Using SIPOC

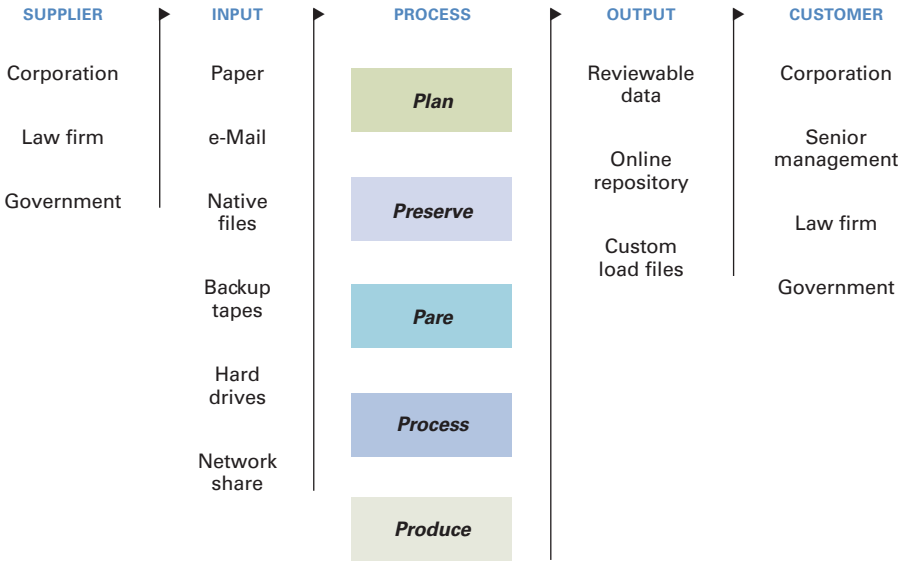
Finally, a key component to helping the team understand the issue(s) at hand is to document the business process associated with the project. Documenting the current business process will help the DMAIC team understand the major components at a high level and identify what major activities come first. It can also help them avoid getting stuck in details.

A SIPOC (pronounced "si-pock") map, as shown on the following page, provides an at-a-glance view of suppliers, inputs, processes, outputs, and customers.

Once this high-level SIPOC map is created, the next step is to break down each process into its component parts. For example, a flowchart could document each step in the planning phase, then the preserving phase, and so on.

Once the project is selected, the charter developed, the VOC translated, and the SIPOC documented, the team is ready to move on to the next DMAIC step: Measure.

EXAMPLE SIPOC FOR DISCOVERY MANAGEMENT ***(High Level)***



MEASURE

Measure is a logical follow-up to Define and serves as a bridge to the next step: Analyze. The Measure step has two main objectives:

- Gather data to validate and quantify the problem/opportunity.
- Begin identifying facts and numbers that offer clues about the causes of the problem.

Some of the key measurement objectives are:

- Determine what to measure and why
- Prepare plans to collect output
- Process and input data
- Construct forms and test data collection procedures
- Refine data collection, if needed
- Refine the DMAIC charter, if needed.

From a discovery management perspective, measurable items might include those outlined in the following table.

MEASURABLE ITEMS FOR DISCOVERY MANAGEMENT

Plan	*Preserve*	*Pare*	*Process*	*Produce*
Budget vs. actual	Number of drives	Set-up time	Data loading time	Creation time for review sets
Number of custodians	Number of servers	Processing time	Metadata extraction time	Required reviewing training time to be proficient
Data size per custodian	Amount of data on drives and servers	Deduplication rates	Number of images rastered (if applicable)	Percentage of time online/available
Approximate page yield per GB	Chain-of-custody defects	Keyword hit rates	File type exceptions	Document decisions per hour
Estimated pages per document	Drives imaged per hour/day	Reviewer training time	Time spent addressing file exceptions	Review call statistics
Estimated time for review	Time spent to image drives	Reviewer decisions per hour/day	Accuracy of document integrity/source information	Time saved with automated production or redaction logs
Estimated number of reviewers	Time spent to restore tapes	Review call statistics		Export quality measures—load file meets requestor's specifications
Percentage breakdown of dataset (% e-mail, % e-docs, etc.)				

Calculating Sigma

Once your data is gathered, calculating the sigma for most processes is relatively easy and can be done with a standard calculator. The key components are definitions and data for the following:

- The measurable "unit" of the deliverable
- The number of customer requirements (or "defect opportunities") for each unit
- The number of customer requirements not met (i.e., defects).

Consider a hypothetical discovery review matter with 40 reviewers and multiple productions. The measurable unit of the deliverable is the "document." The four main customer requirements—correct issue codes, correct privilege calls, correct redactions, and correct bookmarks/annotations of relevant facts—are also the four defect opportunities for each document.

In our hypothetical, 40 people review one production set of 500,000 documents. From that data set we find that 763 documents have incorrect issue codes, 2,455 have incorrect privilege calls, 2,550 have incorrect redactions, and 632 have bookmarks that are not meaningful or clear. To calculate sigma, we take the total number of defects counted, divide by the total number of units, and multiply by the number of defect opportunities:

$$\frac{(763+2{,}455+2{,}550+632)}{500{,}000 \times 4} = \frac{6{,}400}{2{,}000{,}000}$$

…resulting in 0.003200 defects per opportunity.

As mentioned earlier, Six Sigma usually expresses values in terms of defects per million opportunities (DPMO). Accordingly:

$$.003200 \times 1 \text{ million} = 3{,}200 \text{ DPMO}$$

Now, simply look up the 3,200 DPMO in the Sigma Conversion Table (see Appendix on page 30) to determine the sigma level: approximately 4.25 sigma[11] or a 99.7 percent yield.

Cost of Poor Quality

A Six Sigma concept that is helpful in the measurement phase is "cost of poor quality" (COPQ). COPQ measures the dollar impact of problems—both internal (mistakes made that are never visible to the customer, such as additional reviews of documents to correct "produce" or "withhold" decisions) and external (errors found by the customer, such as failure to produce discoverable and nonprivileged documents). Often, Six Sigma teams work with their finance groups to assist with COPQ measurement in current processes.

For any error, defect, or mistake, follow these steps:

1. Count the number of incidents for a period *(e.g., per day or per week).*
2. Determine the labor costs of COPQ *(number of people x hours/days x pay/period).*
3. Determine the material cost *(cost per item used x quantity).*
4. Consider the adverse decision costs *(e.g., cost of sanction, adverse inference, cost of delay, loss of case).*

[11] In the experience of Pivotal Resources, Inc., most service transactional processes that are not previously scrutinized to this DMAIC level start out at 1.5 to 3.5 sigma.

In our example, computing the internal COPQ is relatively simple once the team stops to quantify the labor costs:

COMPUTATION OF INTERNAL COPQ

DEFECT	DEFECT COUNT	MINUTES TO SEARCH/RESOLVE EACH ERROR	TOTAL MINUTES INCURRED	IN HOURS
Issue codes	763	6	4,578	76.30
Privilege calls	2,455	4	9,820	163.67
Redactions	2,550	8	20,400	340.00
Bookmarks	632	6	3,792	63.20
Total	**6,400**		**38,590**	**643.17**
Attorney rate				$200.00
Cost of poor quality (internal)				**$128,633.33**

The team also should document and measure, where applicable, the *external* COPQ. For example, the cost or client impact resulting from incorrect privilege calls may lead to the following questions: What if confidential information were to be turned over inadvertently? What impact would this have on the case or, worse, on overall business continuity? Could the review team work with opposing counsel to retrieve the offending documents? In some litigation matters, as noted in the cases cited in footnote 4, the external COPQ with respect to discovery management can have consequences in the millions, even billions, of dollars.

ANALYZE

In the Analyze phase, the DMAIC team delves into the details of the discovery process, enhances its understanding of problems, and, if all goes as planned, identifies the root cause(s) of the problems. The Analyze phase is like a mystery novel: it can take many paths, twists, and turns. The DMAIC team narrows its search for causes using the analyze *cycle*.

Continuing our example, the analyze cycle begins with the DMAIC team gaining a comprehensive understanding of the document review process, the review time data for various reviewer groups, and review markings. From this information, the team forms a hypothesis as to the cause of incorrect issue codes, privilege calls, redactions, or bookmarks. The team then attempts to support the hypothesis with additional data and other evidence. Often, the first hypothesis turns out to be incorrect. The cycle continues, with the hypothesis refined or rejected until the true root cause is identified and verified with data.

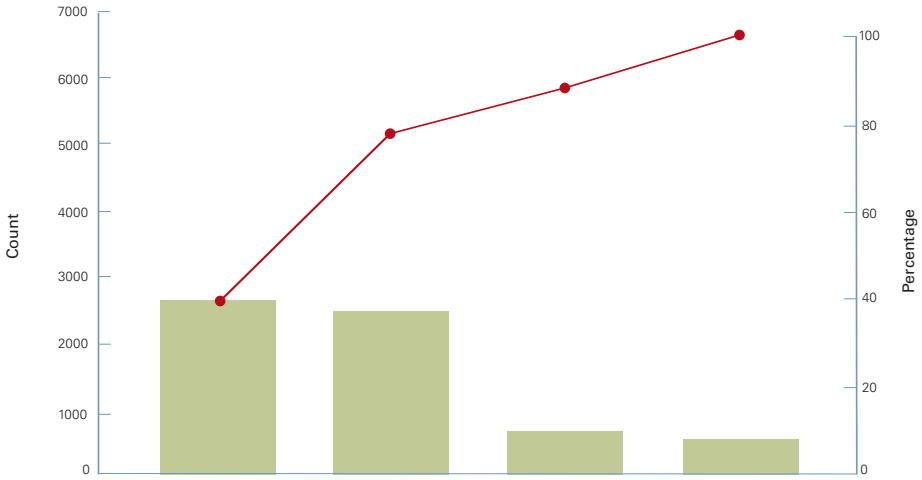
One of the key challenges in the Analyze step is selecting and using the right tools for measuring evidence and data. There are two categories of data analysis tools: (1) visual display tools that help you "see" the data and find clues and (2) statistical tools that test data mathematically in various ways.

Visual Display Tools

Visual display tools are essential to the Analyze phase. These tools include Pareto charts[12], histograms (frequency plots), run (trend) charts, and scatterplots[13]. Many of today's litigation discovery management tools include built-in reporting and simplified data export procedures that can provide easy download and manipulation of potential root-cause data.

Pursuing our example, the DMAIC team digs deeper into the collected data and analyzes the review team's defects by type using a Pareto chart.

DEFECTS BY FREQUENCY



Defect Type	Redactions	Privilege Calls	Issue Codes	Bookmarks
Count	2,550	2,455	763	632
Percentage	39.8	38.4	11.9	9.9
Cum %	39.8	78.2	90.1	100.0

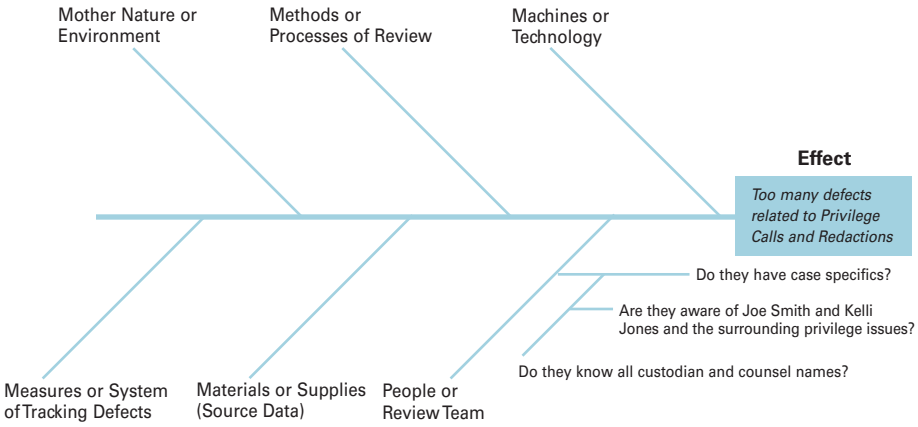
The above Pareto chart displays the review defects in order of most frequent (redactions) to least frequent (bookmarks). The red line tracks the cumulative percentage of each defect as it approaches 100 percent. For example, the cumulative percentage (cum %) of redactions and privilege calls accounts for 78.2 percent (39.8% + 38.4%, respectively) of all the defects in the review set. Accordingly, this gives the DMAIC team an initial indication of where the majority (78.2 percent) of the defects are occurring, thus helping the team to stay focused on what matters most.

[12] The **Pareto chart** is a special type of histogram used to view causes of a problem in order of severity from greatest to smallest. It is a statistical tool that graphically demonstrates the Pareto Principle, or the 80/20 Rule, which is based on Vilfredo Pareto's research—the vital few (20 percent) causes have a greater impact than the trivial many (80 percent). For more information on Pareto charts, see *The Six Sigma Way Team Fieldbook: An Implementation Guide for Process Improvement Teams*, by Peter S. Pander, Robert P. Neuman, and Roland R. Cavanagh, copyright 2002 by The McGraw-Hill Companies, Inc.

[13] A **scatterplot** or **scatter graph** is used in statistics to visually display and compare two or more sets of related quantitative, or numerical, data by displaying only a finite number of points, each having a coordinate on a horizontal and a vertical axis.

Cause-and-Effect "Fishbone" Diagrams

A cause-and-effect, or "fishbone," diagram is a helpful tool that builds on the experience and collective thinking of the DMAIC team. This structured brainstorming technique focuses on generating a list of possible causes of an effect. As shown in the diagram below, the "head" of the fish is the effect and the major "bones" represent primary cause categories.



These categories are often unique to your subject matter, but generally fall into the following six categories:

- Mother Nature or environment
- Methods or processes
- Measurement and information systems
- Machines or equipment
- Materials or supplies (source data)
- People

These categories help the team avoid mistaking symptoms of root causes, such as "deficient training," with actual root causes, such as "the training instructors lack essential case information needed to accurately instruct the review team."

Potential "causes of causes" are then branched out on smaller "bones" to help dig down to the source. Note that the same cause could appear in more than one category. As a rule of thumb, it is helpful to ask "Why?"or "What?" at least four to five times to help get to the root cause.

In our hypothetical example, the line of questioning that would follow the "People or Review Team" bone could proceed as follows:

Why #1: *Why are there too many defects related to privilege calls and redactions?*

Answer: The review team may not have case specifics.

Why #2: *Why doesn't the review team have the case-specific information necessary to make accurate privilege calls and redactions?*

Answer: The review team does not have a list of all custodian and counsel names to track attorney-client communication that could potentially be privileged or require redaction.

Why #3: *What custodian names and counsel names pertain to this matter?*

Answer: A list of 20 custodian and 5 counsel names was provided to the review team.

Why #4: *It the list accurate? What about the surrounding privilege and confidential issues related to Joe Smith and Kelli Jones?*

Answer: We don't see Joe Smith or Kelli Jones on this list of custodian or counsel names. This could be a root cause!

This "why" or "what" exercise is repeated for each potential high-level cause until multiple root causes have been exposed. Continuing in our example, other root causes identified included (1) a potential lack of regular communication among the team, (2) a potential lack of coordinated, consistent training among the review team members, (3) a potentially weak workflow process that relied on manual intervention and double-checking that would often break down or be skipped altogether in time crunches.

After this fishbone exercise, hypotheses are formed and hypothesis statements, a key milestone for the DMAIC team, are documented. For example, in our people and review team "why analysis" above, our hypothesis would be: *We hypothesize that the root cause of too many defects related to privilege calls and redactions derives from the fact that Joe Smith and Kelli Jones were not on the list of custodians or counsel names that was communicated to the review team.*

The team then moves toward investigation and verification of each hypothesis formed by gathering data, observing the cause, applying logic and comparisons, or conducting experiments for this and all other potential root causes.

Statistical Tools

Often, analyzing the process and digging into charts and graphs can provide all that is needed to pinpoint the root cause of a problem. However, the data may be inconclusive or a level of proof beyond what visual tools or brainstorming exercises can offer may be needed. In these cases, DMAIC teams can apply more advanced statistical analysis tools, including:

- *Tests of statistical significance* that look for differences in groups of data. These tests include Chi-square, t-tests, and analysis of variance (ANOVA).
- *Correlation and regression tools*, which are related to scatterplots but can get much more complex, including regression coefficients, simple linear regression, multiple regression, and others. These tools test for the presence, strength, and nature of the links among variables in a process or product.
- *Design of experiments (DOEs)* are a collection of methods for developing and conducting controlled assessments of how a process or a product performs, usually testing two or more characteristics under different conditions.

We do not provide details on these tools for the purposes of this paper; however, comprehensive and formal Six Sigma training typically includes instruction in these types of statistical tools, based on the needs of the project teams.

IMPROVE

After the problem has been defined and data has been measured and analyzed, it is time for the DMAIC team to create innovative ideas for improvement. Following is the road map for the Improve phase:

- Generate solution ideas
- Narrow down the ideas and develop proposed solutions
- Select the best solution
- Plan and test the solution
- Evaluate, tweak, and expand the solution
- Plan a full-scale rollout of the solution

Generating Solution Ideas

Brainstorming can be a useful, engaging tool for generating solution ideas. The purpose of brainstorming is to come up with a list of options for a task or a solution. This list is usually long in the beginning and gets shorter with refinement. Note that one of the challenges of brainstorming is that everybody assumes it is easy. The truth is, effective brainstorming takes work and discipline to be truly creative.

Before the DMAIC team jumps into a brainstorming session, it is important to remember these ground rules:

- Clarify the objective first: "We are brainstorming ideas to address…"
- Record all ideas as given, without editing.
- Everyone should be an active participant.
- No judgment or criticism allowed. There are no "stupid" ideas.
- Hold off discussion until the end of the session (clarification and duplications should be handled after conclusion).
- The more creative/unusual the better!
- Make it fun (order pizza, have whiteboards with lots of pens, etc.).

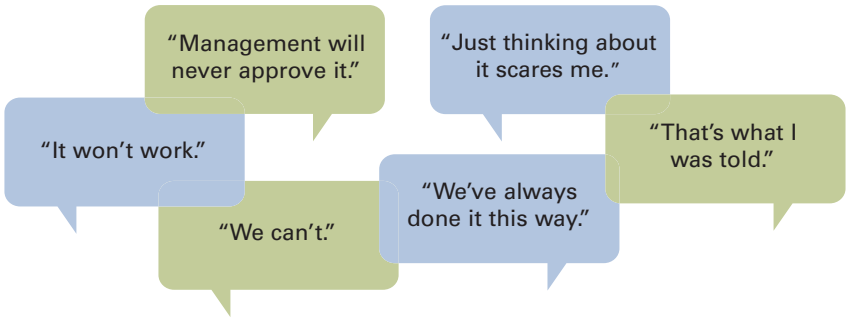
Anti-solutions

Anti-solutions is a brainstorming technique that helps open the team members to seeing things differently. The team first defines the brainstorming objective: "How do we reduce defective privilege calls and redaction errors?" Next, the team defines the opposite objective: "How do we *increase* defective privilege calls and redaction errors?" Team members proceed to brainstorm anti-solutions for this anti-objective and usually have a couple of laughs thinking about all the things that could go wrong. Finally, they mine the anti-solutions for positive ideas they might suggest and expand those ideas into solutions.

Assumption Busting

Beliefs about how things are or how they should be can limit a team's ability to consider innovative, out-of-the-box solutions. Assumption busting is a process that helps identify—and eliminate—preconceptions that inhibit the recognition or creation of viable solutions.

During brainstorming sessions, team members stay on the lookout for ideas (or their own thoughts) that begin with:



Not all assumption-based reactions are bad: they can offer protection from solutions that are impractical or ill advised. However, assumptions can also inhibit the DMAIC team from seriously considering innovative, if unconventional, ideas. Assumption busting provides a system for determining when assumptions are getting in the way.

When an idea sounds good but makes you uncomfortable:

- Identify what premise, rule, or experience is creating discomfort.
- Ask: "Is this premise valid?" "Could I/we be wrong?"
- Consider ways to test the premise. Get expert/outside advice. See how others do it.
- Identify actions you can take to make the new assumptions feasible.
- Make an effort to get used to a new way of seeing things.

Research

Looking outside the team's current, normal environment for inspiration or out-of-the-box approaches also can stimulate solution ideas. This can include looking at company and industry leading practices, company benchmarks, market research, and future or trend analyses. In addition, industry publications, conferences, the Internet, business publications, even mainstream magazines or newspapers can provide information and points of view that prompt new thinking.

Solution Screening

When the team reaches a consensus that idea generation is exhausted, it is time to screen the list and filter for potential solutions. Pivotal Resources suggests these questions that may help narrow the list:

- Does this idea/solution have a clear potential to meet our goal?
- Will it eliminate, or protect the process from, the root cause(s)?
- Will it be acceptable to the process customer?

Further, opportunities may appear for mixing and matching to form solution "packages." Often, the best solutions are a synthesis of several ideas. Finally, the team can develop a weighted criteria matrix that lists what factors and features are key to success, with a numeric ranking to score each alternative.

Improvement Case Example

Building on our continuing example, the data gathered, analyzed, and further developed in the fishbone diagram reveals several possible root causes of defective privilege calls and redactions, which the team brainstorms to identify potential solutions.

Next, the team develops a weighted criteria matrix like the one below that lists potential solutions—alternatives A to D.

WEIGHTED CRITERIA MATRIX

		ALTERNATIVE A		ALTERNATIVE B		ALTERNATIVE C		ALTERNATIVE D	
		Retrain team on case specifics and privilege/ redaction reasons		*Hire additional attorneys to double-check all entries*		*Weekly two-hour team review sessions on Friday to discuss key issues, share case findings, and discuss privilege/redaction issues*		*Leverage technology: deploy software to assist with semantic mapping and auto-categorization*	
Criteria	**Weight**		**S/WS**[a]		**S/WS**		**S/WS**		**S/WS**
Amount of defect reduction likely	10	Possible up-front, fade over time	4/40	Significant	10/100	Significant due to knowledge sharing	8/80	Significant due to "visualization" of data	10/100
Cost, per production, to implement	7	$5,000 investment in hours and travel	8/56	$200 x 500 add'l. hours = $100,000	3/21	$200 x 80 add'l. hours = $16,000	7/49	500,000 docs (approx. 25 gb[b]) @ $3,500 per gb = $87,500	5/35
Time to implement	5	One week	10/50	Four weeks	5/25	One week	10/50	Two weeks	8/40
Positive impact on other parts of review	6	Moderate	4/24	High: other defects may be reduced	9/54	High: other defects or issues may be addressed	9/54	High: other defects or issues may be addressed	10/60
Complexity	8	Low	8/64	Moderate: relatively easy to staff, need to build in workflow processes	7/56	Very low: phone call and/or online meeting	10/80	Moderate: software is relatively easy to use, but training is necessary	7/56
Total Weighted Score			**234**		**256**		**313**		**291**

[a] *Score/Weighted Score: Weight x Score = Weighted Score*
[b] *gb = gigabyte*

There is no rule that says you have to pick only one alternative and abandon the rest. In this case it makes sense that Alternative C (weekly calls) and Alternative D (leverage technology) be seriously considered as potential solutions.

Planning and Testing the Improvement Solution

Now it is time to put the ideas into action and test the hypothesis. The project plan is developed, project management tasks are assigned, and a pilot solution is executed.

In our continuing example, the pilot plan for document review of the next 500,000 documents entails an in-person, two-hour meeting where lead counsel debriefs the review team on relevant case issues (including key custodian information that might lead to

"attorney-client privilege," etc.) and discusses key case topics, deadlines, and milestones. The review team is given an opportunity for questions, answers, and knowledge sharing. The pilot plan also involves implementation of an electronic discovery technology solution to help redefine workflow and improve the review team's quality and efficiency using semantic network mapping[14], auto-categorization, and other document analytic techniques.

During the pilot, the DMAIC team analyzes and monitors the risks and unintended consequences that might result from the implementation of the solutions.

Next, the team tabulates the results of the pilot review and finds that only 125 documents have incorrect issue codes, 110 have incorrect privilege calls, 99 have incorrect redactions, and 115 have bookmarks that are not meaningful or clear. To calculate sigma, we again take the total number of defects counted, divide by the total number of units, and multiply by the number of defect opportunities:

$$\frac{(125+110+99+115)}{500{,}000 \times 4} = \frac{449}{2{,}000{,}000}$$

…resulting in 0.000225 defects per opportunity or 225 DPMO.

Looking up 225 DPMO in the Sigma Conversion Table in the Appendix, we get a new sigma level of approximately 5.0. This represents a sigma increase of 0.75 (5.0 – 4.25 = 0.75) and a yield increase of 0.30 (99.977% – 99.7%). Not much, you say? Take a look at this improvement from a cost-of-poor-quality perspective both internally, with a measurable cost savings of almost $120,000, and externally, with a significant risk reduction of defective documents inadvertently being produced, which could be a multiple of the COPQ savings from internal improvements.

DEFECT	DEFECT COUNT	MINUTES TO SEARCH/RESOLVE EACH ERROR	TOTAL MINUTES INCURRED	IN HOURS
Issue codes	125	6	750	12.50
Privilege calls	110	4	440	7.33
Redactions	99	8	792	13.20
Bookmarks	115	6	690	11.50
Total	**449**		**2,672**	**44.53**
Attorney rate				$200.00
Cost of poor quality (internal)				**$8,906.67**
Original cost of poor quality (internal)				**$128,633.33**
Difference (savings)				**($119,726.67)**
Cost of poor quality (external)				**Reduced**

[14] In the context of discovery management document review, **semantic network mapping** refers to the visualization of data by grouping semantically similar, or "like," documents together based on the co-occurrence of their respective noun or noun phrases in a native review environment. For example, documents related to accounting will likely contain words such as income statement, balance sheet, cash flow, or reconciliation, whereas documents related to personal e-mail may include nouns such as football game, dinner plans, or spouse's name. Accordingly, semantic network mapping technology is able to distinguish between the two categories of documents and group or categorize them separately.

Reporting this back to management, the DMAIC team also considers the return over investment perspective as follows:

RETURN OVER INVESTMENT ANALYSIS

Cost of poor quality savings (internal)	$119,727
Investment cost:	
Weekly meetings	$16,000
Technology	$87,500
	$103,500
Return over investment (internal only)	**16%**

Six Sigma Thresholds

Referencing the goal statement in our charter, we have reduced defects by 93 percent (6,400 DPMO versus 449 DPMO), thereby exceeding our goal of 80 percent defect reduction. Significant efforts would need to be made to ensure 6 sigma, or only 3.4 DPMO. As many DMAIC teams observe, the closer you get to 6 sigma quality, the more difficult it is to identify the root causes and the more radical is the change in process required to achieve such levels. In all likelihood, 5 sigma will be acceptable to management and won't justify further investment in the investigation of root causes: the remaining number of defects is not critical to the overall success of the project (i.e., there is a tolerance for some error in the subjective process of document review). In other situations, the tolerance for error would be so low that further investigation of root causes would be justified. For example, airline companies require 6 sigma (if not higher) performance in airline maintenance. Simply put, from the customer perspective, there's no margin for error—so the airlines must invest heavily in this area.

Given the success of this pilot project, the team moves to integrate the solution into the normal business operations going forward. New processes are developed to anticipate defects and enable preventive actions before defects occur. Questions might include:

- Was the goal met or "in range"?
- Was the desired return on investment acceptable?
- Is the solution sustainable?
- Were the risks manageable?

Finally, a second-wave pilot is considered or a plan is developed for the full-scale rollout, and measures and monitoring are refined with a strong focus on Control. The team can now celebrate a job well done!

CONTROL

An organization is like a rubber band. You can work hard to stretch organizations into all kinds of new and interesting shapes, but as soon as you let go, *snap*…it's back to the old shape. Avoiding that "snap" is the main objective of the Control step. Specific objectives include:

- Tracking the process and results and planning ongoing measurement for future litigation matters
- Recording new ways of working and lessons learned, and developing the documentation
- Preparing revision and response plans as needed
- Maintaining and expanding the gain by considering other parts of the organization and how this innovation could be applied in other contexts
- Considering linkages with other support processes
- Ensuring stakeholder management throughout the project with an eye toward ongoing cultural changes.

The ongoing control plan should focus on a balanced set of factors important to the business and to the "critical to quality" needs of the customers. Control priorities include:

- *Prevention.* Eliminating variation and insulating the process from special-cause disruptions
- *Early detection and response.* Identifying a potential problem or cause prior to it becoming a problem necessitating reactive correction

Planning Ongoing Measurement

With the new solution(s) in place, it is important to update and document the new processes and determine what metrics will best reflect the process performance. Control charts are helpful tools to demonstrate stability and predictability of a process over time. In almost any process, there is variation, even in stable, predictable processes. Control charts will help identify patterns of variation over time and instances of "outliers," which are indicators of sporadic variation or out-of-control situations.

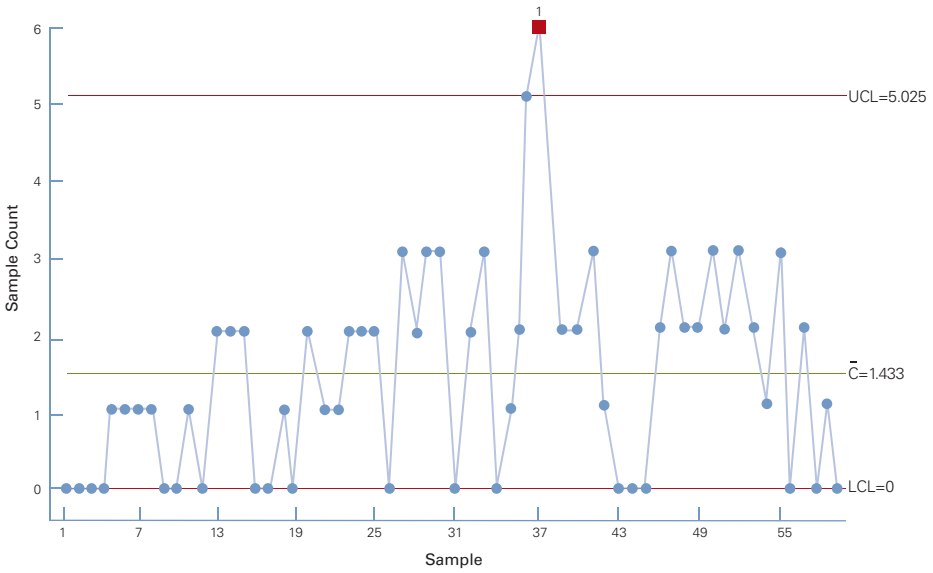
Control Case Example

In our hypothetical example, review defects related to redactions are monitored over a 60-day period. The "C-Chart" on the following page plots the observed defects.[15]

[15] A **C-Chart** is a common format in a service-type business where attribute or discrete data (e.g., defective or not defective) is more prevalent than continuous data (e.g., ranges in oven temperature varying between 350 degrees and 375 degrees). For continuous data, an X-Bar Chart is more appropriate.

With the C-Chart, the objective is to reduce the variation and shift the green centerline (denoted as $\bar{C}$) as close to the lower control limit (LCL) as possible. Since you can't have fewer than zero defects in one day, the LCL is labeled zero. In this case, the team is averaging 1.433 redaction defects per day among a group of 40 reviewers. The team should also look into Day 37 and Day 38 to see why the upper control limit (UCL) was exceeded. Perhaps someone new came on board or some decision criteria were ambiguous.

REDACTION DEFECTS OBSERVED PER DAY ***(Over a 60-Day Period on the XYZ Matter)***



Documentation

Effective documentation must be useable, accessible, and "updateable." To that end, it is critical to make good choices about *what* to document. Criteria might include:

- Activities critical to customer satisfaction
- Complex, multi-step activities
- Processes/tasks that involve many people, especially changing teams of people, such as reviewing attorneys
- Legal or regulatory requirements
- Information about processes that require flexibility.

Easy-to-understand documentation is key to ensuring continuity of your new process. Concise and specific language with minimal, if any, jargon or technical language is always preferable. As a rule of thumb, documentation should be written to an eighth-grade level of readership and should include graphical illustrations. Once written, the documentation should be tested by an objective party. Having a consistent approach to documentation is helpful. At the same time, the team should be prepared to adapt methods according to needs, people, processes, and case requirements.

Six Sigma and Technology

While our hypothetical discovery example is relatively straightforward, real-world problems may be more complex. Hidden root causes masked by years of institutionalized processes and policy can make barrier removal and change difficult. This is particularly true when electronic data is introduced into an environment where paper documents are the paradigm for discovery management control.

Nevertheless, Six Sigma's DMAIC methodology and tools can be highly effective in helping to reduce costs and improve quality. When operating at a 3–4 sigma level, it may be relatively easy to increase sigma by .5 percent or .75 percent by improving on existing processes. DMAIC teams often refer to these types of improvements as "low-hanging fruit."

To successfully break into the 5 and 6 sigma range in discovery management, radical changes in processes and technologies must occur. Top law firms, working with Tier I electronic discovery service providers, are using innovative thinking to reengineer processes, leverage technology where appropriate, and achieve higher sigma scores.

File Preservation/Collection Innovations

Companies involved in litigation or an investigation are often asked to preserve data pertaining to a specific topic. This task often impacts a large number of geographically dispersed data custodians. The task may prove onerous and technically challenging to implement, and it typically requires sending computer professionals to each custodian to manually image hard drives or networks.

One significant innovation being applied to address remote data collection is file preservation program (FPP) technologies. FPPs are sent electronically to remote users along with simple installation and use instructions that enable custodians to identify potentially relevant files on their hard drives and networks. Once all of the relevant files and folders have been selected, the FPP copies and compresses the data and builds a unique, forensically sound container file that can be sent back to the requester in the file's native format.

For each file, the FPP preserves the absolute path—including the original modification, last access, and creation times—and generates a unique digital signature.[16] The tool also includes a comma-separated value (CSV) listing of all files and an audit log that is stored within the container file. Administering a FPP tool takes about one third of the time and cost of physically imaging every hard drive or network share in every location.

[16] MD5 (message-digest algorithm 5) is a widely used cryptographic hash function with a 128-bit hash value.

Document Review Innovations

Legal departments and law firms increasingly are using automated document review techniques to improve the speed and quality of the review process.

Document analytics is the emerging practice of applying algorithms and technology to identify relationships and relevance of documents within a group for purposes of an investigation or discovery. Montgomery Kosma, an attorney with the law firm of Jones Day, works in the firm's Government Regulation, Antitrust, and Competitive Law Practice Group. Mr. Kosma is applying document analytics—specifically, semantic network mapping—to increase his team's document review throughput by a factor of nearly ten. "Prior to our use of document analytic technologies and processes, we estimated our staffing needs based on 250 to 300 document decisions per lawyer per day," says Kosma. "Today, in some cases, our attorneys have achieved 2,000 or more document decisions per day with the same, if not better, accuracy."

Employing document analytic techniques, such as the semantic network map, enables the reviewing attorney to more accurately and rapidly identify relevant information in large electronic document collections. Such techniques, often called "visualizations," also include methods for examining how documents are related by time (e.g., what documents occurred between March and May) and by custodians (who communicates with whom). Advanced visual analysis techniques allow a reviewing attorney to combine visualizations for navigation through the set of electronic documents.

Matter-wide keyword highlighting allows administrators to run searches across the entire collection prior to issuing review sets so that keyword terms of interest are automatically highlighted in both the native and TIFF files for easy review-team identification—thus helping to ensure that nothing gets overlooked. The same applies to matter-wide redactions.

Production history tracking provides a record in the e-discovery database each time a document is produced. When new productions are added, it is often important to know that the document under review has been previously produced in another jurisdiction for the client, especially when court sanctions could be imposed for inconsistent productions.

Workflow enhancements facilitate the management of discovery tasks. For example, an administrator can "carve up" review sets into preassigned queries so that the reviewers can see at login what assignment(s) have been created for them. Meanwhile the administrator, with real-time user-status reporting, can track reviewer productivity statistics as well as assignment completion.

Related documents thread is a feature that facilitates review decisions. When trying to decide whether to produce, withhold, or issue-code a document, it is important for a reviewer to see the entire tree of related documents, including near-duplicates (e-mail threads of a conversation) and related documents (documents with the same "RE:" line in the subject heading) as well as e-mail attachments. Discovery management tools are now capturing all this important information and displaying it in a "related documents" thread. Even the exact duplicate documents that have been suppressed during processing can be displayed if the user desires.

Technology Benefits

The benefits of reassessing processes and applying appropriate technology to aid in the discovery phase of litigation are becoming more widely recognized. *Digital Discovery & e-Evidence* presents the findings of a document review study in an article entitled "Automated Document Review Proves Its Reliability."[17] In the study, 48,000 documents were coded for relevance by traditional "people review" and with the help of e-discovery software. The study found that e-discovery software was 95 percent accurate in identifying relevant documents, whereas "people review" showed only 51 percent accuracy. Using our Sigma Conversion Table, computers operated at a 3.1 sigma, whereas people operated at 1.5 sigma.

[17] Anne Kershaw, November 2005

Conclusion

It's axiomatic in business that you can't manage something unless you can measure it. Applying Six Sigma's DMAIC methodology to your discovery management processes can help your litigation team measure processes and performance, solve problems (not just fix symptoms), and provide an innovative environment that can be the catalyst for breakthrough improvements. It is our hope that other teams or departments will see your "measurable quality improvements" and begin to inquire about applying Six Sigma to other parts of the legal department or to the organization as a whole.

About the Authors

Kelli J. Brooks, Esq., is a principal in KPMG Forensic with KPMG LLP (U.S.). Ms. Brooks focuses on Forensic Technology Services offered by KPMG's Cypress Technology Center (CTEC). She provides corporate law departments and law firms with litigation advisory services, including electronic discovery strategies and discovery management services. A litigation attorney for ten years prior to joining KPMG LLP, Ms. Brooks brings a wealth of experience in the areas of litigation and discovery management. Ms. Brooks can be reached at 714-934-5435 or kjbrooks@kpmg.com.

Carolyn Talasek is a senior consultant with Pivotal Resources, Inc., a Black Belt Six Sigma practitioner, and an experienced workshop leader. Ms. Talasek has helped design and implement effective Six Sigma and process improvement training for many top companies, including Calloway Golf, Cendant Corporation, Cisco Systems, Clorox, Federated Logistics, GE Capital, Sears, Starwood Hotels, and Sun Microsystems. Ms. Talasek can be reached at 480-221-1172 or ctalasek@pivotalresources.com.

About KPMG's Forensic Technology Services

KPMG's Forensic Technology Services (FTS) provides end-to-end electronic discovery services that offer measurable quality and review-team efficiency. In addition to digital evidence collection and computer forensic services, KPMG provides advanced discovery management, data conversion, and data repository services leveraging our Discovery Radar™ Suite of applications and CTEC capabilities and infrastructure.

About Pivotal Resources, Inc.

Pivotal Resources is a global consulting and training firm that focuses on organizational improvement through the implementation of Six Sigma, LEAN, and related methodologies. Using proven strategies, Pivotal Resources can enable diverse organizations to launch and sustain continuous improvement efforts that lead to new levels of customer satisfaction, internal efficiency, and corporate profitability.

Appendix: Sigma Conversion Table (Condensed)

Yield (%)	**DPMO**	**Sigma** (σ)
6.68	933200	0
8.455	915450	0.125
10.56	894400	0.25
13.03	869700	0.375
15.87	841300	0.5
19.08	809200	0.625
22.66	773400	0.75
26.595	734050	0.875
30.85	**691500**	**1**
35.435	645650	1.125
40.13	598700	1.25
45.025	549750	1.375
50	500000	1.5
54.975	450250	1.625
59.87	401300	1.75
64.565	354350	1.875
69.15	**308500**	**2**
73.405	265950	2.125
77.34	226600	2.25
80.92	190800	2.375
84.13	158700	2.5
86.97	130300	2.625
89.44	105600	2.75
91.545	84550	2.875
93.32	**66800**	**3**
94.79	52100	3.125
95.99	40100	3.25
96.96	30400	3.375
97.73	22700	3.5
98.32	16800	3.625
98.78	12200	3.75
99.12	8800	3.875
99.38	**6200**	**4**
99.565	4350	4.125
99.7	3000	4.25
99.795	2050	4.375
99.87	1300	4.5
99.91	900	4.625
99.94	600	4.75
99.96	400	4.875
99.977	**230**	**5**
99.982	180	5.125
99.987	130	5.25
99.992	80	5.375
99.997	30	5.5
99.99767	23.35	5.625
99.99833	16.7	5.75
99.999	10.05	5.875
99.99966	**3.4**	**6**