



HotTopic

FREE!
ROI for IG
HotTopic Mini Web Seminar Series
Register today at www.ama.org



Producing a Good ROI on Information Governance

A Guide to IG Techniques, Tools, Technologies

Open your eyes to sensitive and regulated data

Stored data is a big source of business and compliance risk. It may contain regulated, confidential or sensitive information that must not leak outside your organizational boundaries.

Nuix Sensitive Data Finder applies the power of the patented Nuix Engine to generate Information Transparency™ across the contents of email systems, file shares, archives, databases, and storage systems. Quickly find out what sensitive data you have, where it's stored, and what you can do to protect it.

Start with the facts. Start with Nuix.

Improve Info Governance ROI by Safeguarding Long-Term Electronic Records

Martin Springell

Most institutions tasked with safeguarding permanent and historical records have long recognized that digital content is fragile, requiring a digital preservation strategy that provides a safe mechanism for storing the “bits and bytes,” as well as a way to migrate files to newer formats as older ones become obsolete – ensuring content and records can still be read and used by future generations.

Many organizations are now beginning to realize that the same rigor also needs to be applied to long-term, non-permanent electronic records, especially those with retention periods of 7, 10, 25, 50, or 75 years or more. These types of records are vital to both commercial organizations (for example, financial, legal, and intellectual property) and government and public service institutions (for example, environmental planning, social care, and democratic services).



For most organizations, the consequences of not being able to produce readable versions of these long-term electronic records when required is generally well understood – from fines for non-compliance to the cost of legal challenge, reputational loss, failure to meet legislative mandates, and an inability to leverage corporate knowledge and information for competitive advantage.

Information at Risk

However, the risk of not being able to find, use, or even read these vital long-term electronic records when required is less well understood and is compounded by longer retention times, faster information technology refresh cycles, and the sheer volume and diversity of digital content.

Leading technology analyst Cheryl McKinnon at Forrester Research has gone as far as to say, “The long term retention of digital content means preservation issues must be addressed. Hardware, software and file format obsolescence risks will haunt us if not taken seriously.”

Mitigating Information Risk

Many forward-thinking organizations have already begun to take steps to mitigate and protect themselves against the risk of losing or not being able to read important long-term electronic records, recognizing that simply storing information in an enterprise content or records management system – even as an archive or backup – provides only “bit-level” protection for records and is not a mechanism for ensuring future readability.

Converting electronic records to a format like PDF/A also provides only a partial solution. The PDF/A format might itself become obsolete, but more importantly, it does not accommodate the preservation of interactive and multi-media digital content, such as images, audio/

visual, CAD drawings, spreadsheets, enterprise resource planning systems, web pages, and e-mails.

All this means that building digital preservation seamlessly into the overall information governance lifecycle and implementing policies, processes, and systems that conform with the recognized standards for the long-term preservation of digital content, such as ISO 14721:2012 Open archival information systems (OAIS) – Reference Model and ISO 16363:2012 Audit and certification of trustworthy digital repositories, are critical.

Information Risk Audit

Finally, although not a requirement today, I believe it is only a matter of time before information and risk assessment auditors begin to ask how organizations are conforming to the standards for the preservation and safekeeping of long-term digital records.

You can use this to your advantage to build a stronger return on investment for information governance by clearly articulating the risks (and consequences) of not being able to produce readable and useable versions of important long-term electronic records – and demonstrating how you are mitigating that risk by incorporating standards-based digital preservation into your information lifecycle.

About Preservica

Preservica’s award-winning digital preservation technology is used by leading organizations and institutions around the world to safeguard their vital digital content, collections and electronic records. For more information, go to www.preservica.com and download a white paper at www.preservica.com/resource/electronic-records-preservation.

Martin Springell is product director at Preservica.

Don't Let Perfection Be the Enemy of Good Information Governance

Jake Frazier

For years, the records and information management (RIM) profession has been plagued by an overly theoretical and academic approach to managing organizational data. All too often, RIM managers devote extensive resources to creating lengthy, verbose policy documents that are not widely read, practiced, or enforced – making their quest for perfection their worst enemy.

With new challenges brought forth by *big data* – the continually increasing volume, variety, and velocity of electronic documents – new legal and regulatory hurdles as varied as the Dodd-Frank Act in the United States and “right to be forgotten” laws in Europe, and the demands of e-discovery, much-needed change is afoot in the RIM field.

RIM managers have a wealth of knowledge that can be applied to emerging challenges, and by simply adopting a more practical approach and letting go of intangible processes, real progress can be made toward addressing data overload.

Making IG Work

Most information governance (IG) projects never get off the ground, despite the fact that they are driven by acutely felt pain points. The reason for this is that many IG initiatives take time for benefits to materialize; they rarely offer immediate rewards. However, there is a handful of tactical projects that can help improve employees' day jobs quickly and simultaneously build toward a broader vision. The keys are to forego perfection and start small.

Below are four action-oriented steps RIM managers can take in partnership with legal and other departments to solve pain points and create a sound IG infrastructure.



1. Get rid of legacy data and adjust policy going forward.

A critical first step in IG is to deal with legacy storage by refreshing backups, eliminating storage tapes, and enforcing archiving policy. To remediate legacy back-up tapes as an IG project, legal and compliance must collaborate to take inventory of and address any regulatory and legal hold obligations they contain. This process can also be a forcing function to standardize legal hold policies.

For example, if an organization has 100,000 back-up tapes, perhaps only 100 are subject to current legal holds. At any time, a matter may arise that lays claim to 60,000 of those tapes; but if they are remediated before then as part of an overall and enforced archiving policy, they can't fall under future legal holds.

The key is to take these actions as soon as possible, and put enforceable, sustainable data retention policies in place for all data moving for-

ward. Taking this kind of proactive approach can save an organization millions of dollars in the long run.

2. Bring unstructured data under control.

Every organization contains *unstructured data* (e.g., e-mail, word processing documents) that contains confidential or personally identifiable information that may be subject to privacy laws, such as the Health Insurance Portability and Accountability Act. All organizations are also subject to other regulatory and legal obligations for handling sensitive information. Banks, for example, have a lot of regulated data, such as credit card information, to which many people may have access. And, with serious data breaches becoming commonplace – no industry seems to be immune to them – laws are emerging that prohibit and/or regulate the storage of this type of data.

Data breaches are expensive; the “2014 Cost of Data Breach: Global

Analysis” by Ponemon Institute and IBM found that the average U.S. data breach cost \$5.9 million, or \$200 per breached record.

This issue isn’t going away; it is a critical IG pain point. Scanning file shares for sensitive data, identifying critical information, and getting it under “lock and key” can help mitigate the substantive risks of losing this data, which can also include intellectual property and trade secrets, as

e-discovery. With the extensive awareness most organizations have around e-discovery, a surprising number still uses first-generation tools that have trouble with e-discovery from unstructured data sources, including SharePoint and the cloud, for search and retrieval. Only a handful of organizations – typically the heavy litigators – are on the forefront of emerging technology; in my estimation, about 75% still use basic

team found every personal e-mail archive that was not associated with a current employee or connected to existing legal holds and deleted them. By running that search, the company was able to eliminate 20% of its legacy data before moving to the new system, meaning it would never again have to manage that data, search it for e-discovery, put it under legal hold, or back it up.

Accessing data, understanding it, classifying it, finding business value in it, and taking action on it are really at the heart of what IG is all about. To be effective, IG initiatives must be able to quickly deliver tangible results that can measure impact.

IG projects must be able to show material reduction of risks, successful data migration, faster system operations, improved regulatory compliance, and/or more efficient e-discovery to get the green light from leadership. The IG strategy must have buy-in from the C-suite, or else it will be destined to fall short.

Embrace the Opportunity

The tools needed to implement and sustain the IG process are the same ones used to conduct e-discovery, leaving a lot of opportunity open for records management professionals to help build and sustain programs. If records managers can get out of their comfort zone and focus on practical, tangible programs, they are uniquely positioned to drive them forward and reap the benefits in both day-to-day operations and for the long-term. Always keep in mind that measurable progress in one or two areas of the organization far surpasses any failed attempt at perfection.

About FTI Technology

FTI Technology is the e-discovery and information governance segment within FTI Consulting, a global business advisory firm. FTI Technology’s team includes forensic experts, corporate investigation specialists and technology and e-discovery professionals. For more information, visit www.fitotechnology.com.

Jake Frazier is senior managing director at FTI Consulting. He heads the information governance and compliance practice in FTI’s Technology segment.

IG projects must be able to show material reduction of risks, successful data migration, faster system operations, improved regulatory compliance, and/or more efficient e-discovery to get the green light from leadership.

well as the procedural risks involved with managing a data breach. This is a core building block of overall IG.

3. Modernize message archiving.

E-mail archives are one of the most under-maintained systems within an organization. Most archives are built on aged technology that desperately needs to be updated. Data volumes have grown to a point where archives are bursting at the seams and beginning to fail and crash; this can no longer be ignored. Searching and producing data in the archives for e-discovery purposes can take days or weeks, which can cause serious issues for an organization trying to respond to a regulator.

Newer archiving technology has built-in IG controls and can provide a much better experience for running searches and producing data for e-discovery or other reasons. These controls include preset retention policies and the ability to identify and stop sensitive information from being sent outside the organization.

Replacing or enhancing older archive systems with technology that can build toward broader initiatives provides an easy opportunity to show the value of IG.

4. Move beyond first-generation

imaging of hard drives and pulling e-mail from e-mail archives or servers as their primary processes for e-discovery.

With more savvy regulators and plaintiffs bar than ever before, this is simply no longer acceptable. Most e-discovery requests today extend beyond e-mail and desktop data to include a wide range of data sources, such as file shares and collaborative platforms, which often contain information critical to a case.

From a tactical standpoint, updating e-discovery software can enable the legal team to deal with matters that are on their desks today. But, it can also build toward better IG.

IG Done Right

A recent storage migration at a credit card company serves as one example of an IG effort done right. This client was switching from one archiving vendor to another and planned to migrate a petabyte of data as part of the new agreement.

Stakeholders from legal and IT analyzed the data stores before the migration took place, identifying what data needed to be saved and deleting the rest with the intention of optimizing the new storage from the start.

This process was very simple. The

Untangle the Ear Buds



Free Web Seminars from HotTopic

Listen to the spring/summer series on
How to Produce a Good ROI on Information Governance
with presentations from our sponsors:



www.arma.org/r1/publications/hot-topic

Ensuring Information Governance ROI Should Start with the Facts

Julie Colgan

The Information Governance Initiative's (IGI) 2014 annual report estimated that the average small to medium-sized business is spending a total of \$1.3 million across four information governance (IG) projects. The average enterprise is spending around \$12 million across six IG strategies. This is a lot of money to spend without a clear understanding of what the business gets in return or whether these are even the right initiatives to be focusing on. (Register at www.igiinitiative.com to access the report.)

The IGI annual report also highlights that IG – and the product market that supports it – is broad. The report listed the five most popular IG products as “records and information management, information security, compliance, e-discovery, and data storage and archiving.” While all of these capabilities certainly do support governance of enterprise information, they serve vastly different purposes. How can an organization know which of these tools is the right tool for them, when to buy it, how much to buy, and what the return on investment (ROI) measurements and expectations should be?

Something Is Missing

What is missing from the list above is technology that enables organizations to get to the facts in and about their data; to let the data itself help drive priorities and investments. If you can learn what information you have, understand its importance, and classify it, then and only then can you make smart investments in downstream information technologies such as lifecycle and security management.

If enterprises had this kind of technology back in the day, I suspect e-mail archives may not have been so wildly popular. If IT could have evaluated e-mail on its content and context – at scale – it could have enabled more appropriate and customized approaches to how e-mail stor-

age and costs could be contained without making additional purchases or hamstringing the business by instituting draconian mailbox size limits that only resulted in the creation of e-mail black markets.

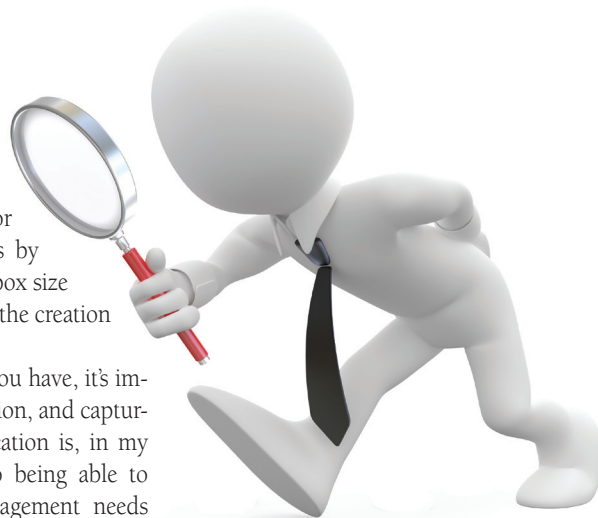
Knowing what data you have, its importance to your organization, and capturing that value via classification is, in my opinion, a prerequisite to being able to articulate your data management needs clearly, including defining anticipated ROI.

For example, an organization wants to start using an enterprise content or document management system for file storage, rather than use simple file shares, to better enable enterprise information sharing and collaboration. However, it is common to find that file shares have a high amount of duplication and old and transitory content that doesn't warrant retention. Having that insight upfront will enable the organization to move only what is truly valuable to the organization thereby saving itself the time and effort of moving things of no value, as well as saving money since storage attached to such systems is often more expensive than simple file storage.

These kinds of technologies also enable organizations to be more proactive in knowing about, understanding the importance of, and making purposeful and appropriate decisions about how best to manage confidential and sensitive data like intellectual property (e.g., trade secrets, proprietary recipes, manuscripts) and private data (e.g. personal IDs, payment card data). In fact if this practice were more common, it is likely that the high-profile data breaches that have been in the news lately may have been far less dramatic and impacted far fewer people.

These Technologies Needed

I propose that, while there certainly is a need for a multitude of specialized



technologies to solve the broad needs of full-scale information governance, specific technologies fall into broader categories than has been previously defined:

- Discovery, Analysis and Classification
- Lifecycle and Workflow
- Security/Protection
- Storage Management

The class of technologies that enable discovery, analysis, and classification are those that uncover and highlight the true data management needs of the organization and will therefore optimize the timing, nature, and level of investment in the other classes of technologies. The more detail, depth, breadth, scale, and precision “discovery, analysis and classification” technologies can expose, the higher the value it will have to the organization across a multitude of IG activities such as information management, e-discovery, and compliance.

There's no need to guess anymore – start with the facts.

About Nuix

Organizations turn to Nuix for fast, accurate help with digital investigation, cybersecurity, e-discovery, information governance, e-mail migration, privacy and more. For more information, visit www.nuix.com.

Julie Colgan, IGP, CRM, is head of information governance solutions at Nuix and chair of the ARMA International Board of Directors.

Information Governance ROI = Reduce Risk, Organize Infrastructure, Increase Value

Kevin Ogradnik

allowing us to think of ROI in slightly different terms:

R = Reduce risk

O = Organize infrastructure

I = Increase value

Thinking of the ROI for IG from this perspective helps make a strong case for implementing a solid CIGP and provides a foundation for it to be successful.

R=Reduce Risk

Every organization has a responsibility to protect their employee, customer, vendor, and other business information. This means protecting the information's confidentiality, integrity, and availability. For many, it also includes complying with legal and regulatory standards for the information. The cost of non-compliance or of ineffective security and management is heavy – as witnessed by all the recent data breaches that have grabbed the headlines. But the cost is not only financial, it is also reputational.

Information risk can range from natural disasters to system outages, theft, fraud, unauthorized disclosure, and inadequate retention periods and policies. Reducing the risk associated with information has to be a strategic concern that aligns with business goals and capitalizes on opportunities rather than focusing on just loss avoidance.

A CIGP helps shift responsibility from line managers focused on an ad hoc approach to strategic-level stakeholders with a documented strategy and accountability. Its benefits include improved decision making, minimized operational disruptions, compliance with applicable laws and regulations, and the ability to react quickly to potentially costly remediation strategies.

A CIGP is one of the most effective



The world of business – or at least the world of the information professional – is all abuzz about *information governance* (IG), which has a variety of definitions but is defined by Gartner as “... the specification of decision rights and an accountability framework to encourage desirable behavior in the valuation, creation, storage, use, archival and deletion of information. It includes the processes, roles, standards and metrics that ensure the effective and efficient use of information in enabling an organization to achieve its goals.”

However IG is defined, everyone is talking about how it can be used to manage the challenges that come from the rapid pace of electronic information growth. These challenges not only have a major negative impact on user productivity, but also on the time, costs, and liability associated with increased storage needs, difficulty in complying with regulations, and the ability to find information when it is needed for litigation and e-discovery requests.

Governing the Info Lifecycle

More simply put than Gartner's defi-

nition, IG is a set of interdisciplinary policies and procedures used to regulate an organization's electronic assets from creation to disposal, or though the information lifecycle. To quote Sherpa Software's vice president of strategy, Rick Wilson, “The ultimate goal of information governance is to recognize that information generated by day-to-day operations of an organization is a valuable corporate asset that must be managed and disposed of in a responsible fashion.”

Many companies, even those in heavily regulated or litigious environments, are wading very slowly into their data pool to implement what Sherpa refers to as a corporate information governance program (CIGP®). Many are afraid to dive in head first without a realistic view of how deep the commitment may run and what return on investment (ROI) they can expect.

Measuring the ROI of IG is a difficult task because of the long-time mindset that managing information is about managing risk and costs. It is very difficult to flip that discussion to focus on the *value* of managing information. IG elevates managing, protecting, monitoring, and disposing of information to a more strategic level, al-

ways to reduce exposure to operational risk. This return doesn't necessarily show up on the top or bottom lines, but the value is immeasurable.

O=Organize Infrastructure

If reducing risk and increasing value are the keys to your CIGP's ROI, organizing your infrastructure is the glue that holds it all together. Three main actions are critical to a successful CIGP: organizing your IG committee, organizing your data, and organizing your IG technology. Consider these key components the foundation of your IG strategy.

Organizing the IG Committee

To be successful, a CIGP should be an enterprise-wide initiative that supports the organization's overall business objectives and is endorsed by senior management. Since IG will ultimately touch every area of a business, it's important to have an IG committee responsible for its implementation and ongoing management and auditing.

Rick Wilson describes an ideal IG committee: "IG committee members should represent a cross-section of the organization in order to bring diverse expertise and knowledge to the project. Typically, the committee will be represented by various departments that have direct knowledge of, and potential responsibility for, handling your organization's internal and external data requirements. This may also include regulatory requirements."

He explained, "Most IG committees have representation from the executive team, compliance, IT, HR, legal, records, and/or security. The IG committee members should know where the organization's data is kept, what information needs to be stored, how long it should be stored, what information should be deleted, when it should be deleted, and how information is accessed and moved within the organization."

He added this advice: "Treat your committee as a group of trusted advisors; they will have the knowledge to help you identify which areas of the business can benefit most from an information governance project, what the degree of difficulty will be to implement that initiative, and

how best to socialize the project within each segment of the business."

Organizing the Data

The IG committee members will be instrumental in organizing your organization's electronic information. One of the most difficult aspects of information governance is knowing what information exists, where it is stored, and how to classify it. Therefore, creating an information inventory is the next critical step to ensuring your ROI. This is a difficult and time-consuming exercise, but once done, allows

Three main actions are critical to a successful CIGP: organizing your IG committee, organizing your data, and organizing your IG technology. Consider these key components the foundation of your IG strategy.

you not only to classify it by content, but also to apply retention periods to it.

Implementing retention policies will also help eliminate redundant, outdated, and trivial information (ROT), which is information a company keeps but doesn't need, such as e-mail conversations between coworkers, outdated contact information, and duplicate information.

According to AIIM's May 2014 Industry Watch report "Automating Information Governance," as much as 80% of electronically stored information is ROT. That's a lot of non-essential information taking up vital storage space, increasing information risk and storage costs, as well as wasting time and resources. ROT can easily be found and removed, as long as you know what to look for and where.

Organizing the Technology

Relying on users alone to execute written policies is simply not a defensible practice. Implementing automated policy enforcement tools is the best way to overcome the hurdles associated with this approach.

So, where do you start? First, know your options: there are solutions for organizations of all shapes and sizes that will locate and inventory unstructured data throughout the organization. In fact, there

are many enterprise platforms available to manage the entire IG process as well as many a la cart options for specific specialization. These will help you visually assess core areas of IG concern, including ROT, intellectual property and risk, and data inventory. From there, your organization will know which data to eliminate by applying the correct retention policies.

I=Increase Value

The key to extracting value from information is closely tied to its availability to the business users who need it to op-

erate the organization. Information access and integrity are crucial, and they require your organization to have and enforce specific information policies, processes, and protocols.

As the foundation for extracting value, information must be classified and stored in the appropriate location. If the lines of business – who need data to create revenue and profit – cannot find and access the proper data when and how they need it, the data becomes useless.

Organizations that treat their data as an asset rather than a liability are able to derive much more value from their data and will therefore see a much greater ROI from their CIGP.

About Sherpa

Sherpa Software is a leading provider of technology-driven information governance solutions, including software, services, and support that address information management, regulatory compliance, electronic discovery, PST management, e-mail archiving, and more. For more information, visit www.sherpasoftware.com.

Kevin Ogrodnik, president of Sherpa Software, leads the company's corporate strategy, drives product direction and positioning, oversees company operations, and generates awareness for the company.

Document Classification and Metadata Guide to Higher ROI

Greg Council

Most organizations struggle with process inefficiencies and the inability to locate and use documents within the normal course of business. These problems can often be traced to inconsistent rules governing where documents are stored and how they are described or inconsistent enforcement of information management policies. The fact is that organizations cannot rely solely on either technology or staff to completely solve the problem. Without a simple, manageable place to start, how can any organization address these challenges?

Unless an organization is familiar with the latest technologies and how to implement them, the typical avenue to solve the problem is through brute force: hire new staff or use existing staff to manually review each document to assign it to a category, catalog it, and then update the new document control system. The cost of this path is often too high, either in workload or additional costs.

The good news is that there are proven solutions and software that can automate this task at a fraction of the labor and time costs of alternative manual review and tagging. To achieve a technology-driven classification and metadata project requires five major steps:

Step One: Take an inventory of all document assets. As stated earlier, one way is the brute force method. Have staff manually review every file. There is also software that can automate the clustering of documents by likeness in much the same way a human would. The actual capability is called document clustering, and it can automate grouping by visual elements or content, or both, to dramatically reduce the amount of work required. Import the volume of documents to be grouped and the software automatically sifts through them, evaluates each document, and places them into clusters.



Step Two: Identify each document type. For this task, it's key to have staff subject matter experts (SMEs) familiar with the types of documents used in business. Manually assign these documents to their respective types, or take a small subset of each type (identified by the SMEs) and use this subset to automatically create rules using classification software. An automated approach allows you to forego the time-consuming task of analyzing each type of document and variants in order to discover rules that can be used and applied automatically. After the initial rules-creation process, an analyst can review results and make adjustments – it doesn't have to be a “black-box” approach.

Step Three: Identify your metadata. After establishing a comprehensive view of all document types, the next step is to figure out what key data about each docu-

ment type is useful for management. It's important to think beyond activities, such as document search and retrieval, to include governance-related data that facilitates access control and privacy policies. Document types and metadata provide the basis of a true document taxonomy.

Step Four: Implement metadata rules. Each document type needs metadata rules applied to it so when it is automatically assigned a class, data can either be extracted or applied to the document. Once again, software can automate this task by using the document class assignment to automatically apply metadata rules without any need for manual tagging.

Step Five: Start the classification and metadata assignment process. Simply import the volumes of documents and the software automatically assigns documents to their respective classes, extracting and assigning metadata and then exporting it to your destination of choice. Documents that cannot be classified appropriately are handled by an automated workflow. Automation of the quality control process is also an important time-saving activity, as all documents go through the same process and there is a complete audit record.

This step-by-step guide to how to use software and automation will cut-down on costs, time, and complexity of any document classification and metadata project for efficient information governance.

About Parascript

Parascript is the industry leader in economically extracting meaningful information from a wide range of images, handwritten documents to unstructured forms so organizations increase efficiency and productivity at reduced cost. For more information, visit www.parascript.com.

Greg Council is the vice president of marketing and product management at Parascript.

IS INFORMATION YOUR ALLY OR YOUR ENEMY?

**NEXT
LEVEL™**



information
governance
assessment

Your business data is doubling by the year. And all this new data can either help you or hurt you. Find out what it's doing for your company with the Next Level Information Governance Assessment. You'll discover areas of strength and opportunities for improvement. In the end, you will be empowered to increase organizational transparency and data integrity while decreasing risk.

Start turning information into an asset by visiting arma.org/nextlevel.

ADVANCED DOCUMENT CLASSIFICATION

FOR INFORMATION GOVERNANCE



ANY DOCUMENT



ANY DATA



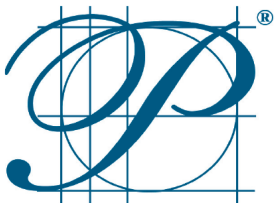
ANY SOURCE

Parascript® Document Classification combines an easy-to-use analysis tool with a powerful production SDK that enables organizations to efficiently and quickly organize document-based information to:

- Ensure Compliance
- Support Discovery
- Organize Documents for Migration
- Reduce ROT (Redundancy, Obsolete, Trivial) Data

Visit us online at **Parascript** to find out more about how our technology supports your clients' business needs.

Latest Step-by-step
Guide to Document
Classification &
Metadata



PARASCRIPT®

888.225.0169



Request a Demo

info@parascript.com

www.parascript.com