



Electronic Content & Record Management Best Practices

*Guidelines for Proactive Ediscovery: Improving Data
Accessibility and Simplifying Compliance
with Federal Rules of Civil Procedure*

by Tom Leh, Managing Partner, Avacuna LLC

INTRODUCTION

The disposition of digital assets in the modern enterprise is undergoing a fundamental and revolutionary transformation. For decades, businesses have focused primarily on how to compile valuable information within the organization and share it between functional units. Convergent regulatory, technological and user convenience requirements are compelling companies to consider how best to confront emerging information management challenges.

Although frequently lagging the pace of technological innovation, regulatory attempts at classification are inevitable and already in progress. In December 2006, the Federal Rules of Civil Procedure (FRCP) were amended to include an emerging information category identified as electronically stored information (ESI). The FRCP's Rule 34(a) defines ESI as "...including writings, drawings, graphs, charts, photographs, sound recordings, images, and other data or data compilations stored in any medium from which information can be obtained..." In addition to horizontally applicable governance obligations that affect all businesses, sector-specific changes loom over the horizon for several verticals. Financial institutions will need to consider their planned response to the Federal Deposit Insurance Corporation's amended capital adequacy rules for Basel II, which advise that, "...a bank must retain data in an electronic format that allows timely retrieval for analysis, reporting, and disclosure purposes." (72 Fed. Reg. 210). Given the absence of material comments on this broadly defined data requirement, it is reasonable to expect additional refinement going forward as unforeseen legal and fiscal reporting issues evolve.

Within the context of emerging technologies, the distinction between customers and employees has blurred nearly to the point of irrelevance: both 'internal' and 'external' users have developed expectations of convenience for how they utilize nascent communication tools, and can be classified simply as 'users'. The world of modern data and/or ESI includes an overwhelming array of options for users including mobility devices and emerging applications with messaging, VoIP, peer-to-peer, web content, chat, blog, RSS and wikis that are inherently insecure – yet nearly impossible to regulate using policy-based software. Since these types of utilities offer irresistible collaborative benefits, their inevitable adoption in the business domain requires companies to acknowledge the risks and liabilities associated with the absence of sound digital asset and ediscovery archiving strategies. A mitigation blueprint for ediscovery governance best practices should incorporate the following elements.

1. Understand Data's Multiple Personalities

When the definition for ESI is applied, it becomes apparent that business data can exist in a panoply of forms. Traditional physical documents will include reports, claims, forms, books, maps, requisitions and anything recorded on paper. Contemporary electronic documents may include data originating from word processing, spreadsheet, email, messaging, multimedia, CRM, ERP, IVR, facsimile, web and telephony applications.

In the early phase of any proactive ediscovery plan, an organization should inventory the electronic data formats currently in use by its business units. Since all forms of ESI are discoverable under FRCP Rule 34(b) and requesting parties may ask for specific and/or native file formats, companies that apply 'scan only' or 'print only' practices for electronic file storage are at risk. An archiving process that preserves native file formats is recommended.

Where feasible, depending upon the company's size and scope, it is helpful to establish basic metadata (a.k.a. embedded data or 'data about data') and data classification policies for newly created documents that prioritize subject and content over format. For example, Microsoft Word's document properties options allow for author, title, version numbers, internal reference numbers and change tracking metadata entries. While exemptions (aka "safe harbor") exist for the routine good faith disposal of computer systems data, establishing a data classification culture helps to reduce superfluous data creation as well as aid archiving efforts.

2. Recognize Data Generation Sources

Organizations that embrace a holistic interpretation of data stand a better chance of progress with cost-effective ediscovery compliance planning and execution. Data can originate from a computer, fax, telephone, mobile appliance or a human hand with a piece of paper and a pen. Pervasive tools such as web content, email, instant messaging, collaborative applications, knowledge management and social software are utilized both inside and outside the office and catalyze manipulation of data on a massive scale.

The assumption that every form of data must be evaluated and classified is one intuitively perceived in most companies. What is often more difficult to reconcile is the countervailing threat to individual privacy posed by the classification policy ultimately adopted. As event correlation devices such as wireless LAN access points, GPS, PDAs and cell phones continue to evolve and generate data in business environments, information managers must cautiously consider rule-based access processes. The context, place and manner of some generated data must be carefully balanced such that the archiving of personally identifying information does not circumvent Ediscovery privacy objectives.

After an initial data source assessment, the organization's topology needs to be documented. Ediscovery program stakeholders should bear in mind two significant FRCP rules during documentation and policy development: native file formats per Section 34(b) and the implied linkage between metadata and accessibility per Section 26(b)(2). Decisions on native file formats should be evaluated and incorporated into the policy process with the understanding that requesting parties in an Ediscovery action can solicit "reasonably accessible" information. In the absence of a showing that an undue cost burden is present, the party from whom discovery is sought will be expected to provide native file formats. Since metadata procedures are neither difficult nor burdensome to implement, it is reasonable to anticipate that metadata and hidden classification data will be construed as accessible. Given this condition, smart businesses will transform this obligation into an advantage, and leverage metadata standardization into the data management process preemptively to simplify classification and retention.

3. Use Common Sense to Determine Safe Harbor Applicability

Under Rule 37(f) of the FRCP, a "safe harbor" for automated deletion of ESI exists as part of any party's "routine, good faith operation". Absent the existence of an explicit data retention guideline mandating storage duration periods (e.g., health care, life sciences, financial services, SOX, etc.), it is acceptable to establish deletion policies. Some generated ESI will have no relevance for any legal proceeding, and it would be unreasonable to mandate the storage of electronic content generated by every human resource within an organization. Additionally, some emerging messaging technologies are not easily captured and archived, and not all organizations have tools to deploy for this purpose. Examples of

inaccessible information that may expect reasonable protection from an ediscovery request may include:

- Data fragments or residual data from deleted disk drives and/or data wiping sequences
- Duplicative data in a secondary file format already accessible in its native or source format
- Back-up data intended for use in disaster recovery or business continuity circumstances and defined as such by organizational policies
- Obsolete systems and/or legacy data

4. Recognize Content Management Solution Limitations

Information and knowledge management professionals will need to reconcile their ediscovery objectives with trends that seek to disrupt and impair a perfectly structured content environment. On the one hand, businesses should favor extensibility and federated records management capabilities in their analysis of vendor solutions. Reacting to the market's frustrations and expectations for solutions in this area, a few first tier suppliers have integrated respectable record management, archiving, imaging, security, document management, digital asset management and web content management features into their offerings.

At the same time, several unavoidable trends driven by both individual preferences and business efficiency will force compromises as to the degree of centralized control over data. More users are adopting thin-client mobility, unified communications and collaborative networking tools while distributing email content to ever-expanding numbers of recipients. The ease and pervasiveness of email usage contributes to decentralized file management and less reliable versioning controls. From a business efficiency perspective, increased adoption of Software-as-a-Service (SaaS), open source and hosted applications to satisfy flexibility and functionality demand will feed overall systems growth. As more systems are introduced and/or replaced by alternative options at ever shortening intervals in any organization, complexity will increase geometrically.

Improving efficiency with the management and archiving of existing data records and documents is a priority. It is also important to examine how employees use newer office productivity technologies and favor solutions that extend some degree of federated control and retention policy management over both legacy content sources and the most pervasively adopted workplace tools. While the content management discipline has recently focused increasing emphasis on analyzing how content is put into context and consumed by people and business processes, evaluators of application platforms will be best served by not expecting a single solution to satisfy every strategic objective in toto. At this point in content management's evolution, facets such as personalization, publishing, digital asset management and web content management remain too dynamic for any one platform to tightly consolidate. Focus on business process and record management fundamentals, and embrace the inevitability of a multi-vendor environment.

5. Simplify Data Classification and Retention with Metadata

One of the most critical aspects of ESI Ediscovery revolves around metadata, its use in data classification and the significance of establishing policies governing it in different scenarios. Few companies are sufficiently prepared for the vital role metadata will play in the future of

intelligent content management, and the policy governing its use is a two-sided coin. Metadata represents a complex challenge as it presents both efficiency-based storage benefits and potential liabilities vis-à-vis its innate power to jeopardize confidentiality.

Metadata can be leveraged by a system to find and display content easily and consistently. Companies must solicit the development of policies and procedures for the creation, preservation and management of native file formats and planning metadata inclusion is a natural extension to any strategic Ediscovery plan. Through the inclusion of a few defined parameters such as a document's title, summary description, version number, internal reference number, forwarding history and/or change tracking options, a granular and secure content management environment can be readily established. Data categorization of this nature has become increasingly vital for effective regulatory compliance and data retention management efficiency in an era when data capacity is doubling every ten months.

While generally beneficial, metadata is also a sensitive information asset burdened by potential confidentiality and privacy considerations. There are instances in which deletion is recommended for either practical operational reasons or ethical obligations within a legal context. Automated metadata deletion procedures are protected under the FRCP "safe harbor" exclusion when undertaken as part of good faith routine operations. In Florida, an attorney sending an electronic document is expected to ensure the confidentiality of all information contained in the document *including* metadata; other state bars impose similar expectations. While the Florida Bar notes this guidance is not intended to govern metadata in discovery document applications, it exposes the perils of oversimplifying metadata policy planning and emphasizes the need for collaborative approaches between line of business, IT and legal stakeholders. Each group must consider strategic business objectives from their equally relevant, yet contrasting perspectives, to refine and approve metadata policies that satisfy both internal and external data scenarios.

SUMMARY

Nearly every facet of Ediscovery governance is dynamically evolving. Legal precedents, compliance programs, ESI record management, automation tools and metadata management controls are all subject to ongoing modification. Early program implementation progress can be diminished by an absence of executive management evangelism for Ediscovery's growing importance. For this reason, organizations should view the process as iterative and expect several revisions prior to gaining confidence in a full production system. Interdepartmental cooperation incorporating line of business, IT and legal perspectives as early as possible will accelerate process refinement and assist with socializing Ediscovery hygiene throughout the organization.

Adherence to the following principles will help those organizations committed to successfully maneuvering through Ediscovery planning bottlenecks:

- Develop an information taxonomy to classify newly created ESI
- Define record retention processes only after thoroughly evaluating the 'chain of custody' for a given document and its metadata controls and capabilities
- Establish metadata guidelines for both internal and external scenarios and enforce both consistently
- Finalize processes reinforced by reasonable business justifications to maximize "safe harbor" protection: avoid policies calling for excessive deletion of metadata or overlooking it altogether

Once an organization's information management plan is finalized, employees must be trained to socialize awareness of laws and company policies governing privacy, confidentiality and metadata. No expectations of privacy should be established for the use of portable data devices, email, unified communications, computers or network storage. In optimal environments, employees should undergo periodic evaluations for their understanding of Ediscovery compliance procedures and execute written agreements with the company to bind them to confidentiality obligations.

REFERENCES:

Bace, J. (2007). The New Federal Rules of Civil Procedure: IT Opportunity, Legal Albatross, or Vendor Hype & FUD Opportunity. *The New York State Forum Seminar: Document Management - The 3 R's - Requirements, Retention and Recall*, Albany, February 27, 2007.

Center for Digital Government. (2007). *For the Record: Maintaining Trust in Public Institutions by Ensuring the Integrity of the Digital Public Record*. Folsom: e.Republic, Inc.

CMS Watch. (2007). Building Blocks of Information Access: Information Architecture, Content Management, and Search. *EPA Environmental Information Symposium, Lexington, November 15, 2007*.

Federal Register. (2007). *Risk-Based Capital Standards: Advanced Capital Adequacy Framework—Basel II; Final Rule, Federal Register 72:235, December 7, 2007*.

Frappaolo, C and Keldsen D. (2007). Content Security at the Fulcrum of Innovation and Risk. *AIIM - The ECM Association*. Silver Spring: AIIM.

Harnish, M. (2007). E-Discovery: Do you Know Where Your Client's Data is (or Where it's Been)? *InfoTech Update - Newsletter of the AICPA Information Technology Section, September/October, 2007*.

Kennedy, D. (2007). *FRCP and Metadata: Avoiding the Lurking e-Discovery Disaster*. San Francisco: Workshare, Inc.

Maynard, K. (2007). *Ethical Obligations Arising in Electronic Discovery, November 2007*. Charlotte: Robinson, Bradshaw & Hinson P.A.

McNabb, K. (2007). *The Forrester Wave: Enterprise Content Management Suites, Q4 2007*. Cambridge: Forrester Research, Inc.

Mullins, C. (2006). *The Impact of Regulatory Compliance on Data Management*. Sugar Land: NEON Enterprise Software, Inc.

Osterman Research. (2007). *Why Your Organization Needs to Focus on Outbound Content - An Osterman Research White Paper*. Black Diamond: Osterman Research, Inc.

Rasmussen, C. (2007). Content and Record Management in the Information Age. *Financial Services Technology Forum 2007, Toronto, October 24, 2007*.

Rothstein, B and Hedges, R and Wiggins, E. (2007). *Managing Discovery of Electronic Information: A Pocket Guide for Judges*. Federal Judicial Center.

The Florida Bar. (2006). *Professional Ethics of the Florida Bar, Opinion 06-2, September 15, 2006*. Tallahassee: The Florida Bar.

The Sedona Conference. (2007). *The Sedona Conference Glossary: E-Discovery & Digital Information Management*. 2nd Ed. Sedona: The Sedona Conference.

United States. (2007). *Federal Rules of Civil Procedure, December 1, 2007*. Washington: GPO.

ABOUT AVACUNA LLC

Avacuna LLC is a provider of IT compliance, risk assessment and electronic content management consulting services. Avacuna's practice areas combine knowledge of information lifecycle management, general computer controls, internal audits, ediscovery strategies and regulatory analysis. Avacuna exposes the transactional value creation potential of information systems governance frameworks and processes to transform underutilized knowledge capital into competitive advantages and improved valuations for clients and their stakeholders.

Avacuna LLC

www.avacuna.com
info@avacuna.com

Boca Raton, FL

954.719.5126