

# Managing Electronic Records

May 19, 2009

## Metadata Management: The New ERM Frontier?



Julie Gable



# Why the “sudden” interest?

- Recent Court Case
- Sedona Conference Changes
- Metadata Standards
- ECM and Governance Solutions



# Aguilar v. Immigration & Customs Enforcement Div.

- Class action alleging unlawful searches and seizures of plaintiffs' homes
- Court compelled production of metadata as part of e-discovery

2008 WL 5062700 (S.D.N.Y. Nov. 21, 2008)

# Aguilar v. ICE

- Court identifies 3 kinds of metadata:
  - Substantive – application based; may contain modifications, edits, comments
  - System-based - includes author, date, time of creation, date of modification
  - Embedded – text, numbers, content directly input but not visible on output, e.g., spreadsheet formulas, hyperlinks

# Aguilar v. ICE

- Court applies FRCP Rule 26(b)(2)(C)
  - Metadata is discoverable if
    - It is relevant to the claim or defense
    - It is not privileged.
  - Metadata subject to discovery balancing test
    - Court weighs probative value of metadata against burden to produce it



# Aguilar v. ICE

- Court consults Sedona Principles
  - When selecting form of production (per FRCP 34) two primary considerations:
    - Is metadata relevant?
    - Does it enhance the utility of documents, including the ability to search them?

# Sedona Principle 12

- 2004: “Unless material to resolving dispute, there is no obligation to preserve and produce metadata absent agreement of parties or order of the court.”
- 2007: The form of production should take into account “the need to produce reasonably accessible metadata that will enable the receiving party to have the same ability to access, search and display information as the producing party where appropriate or necessary in light of the nature of the information and the needs of the case.”

The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production, BNA Publications, 2007.

# Aguilar v. ICE

- Court addressed case law
  - “Courts generally have ordered production of metadata when it is sought in initial document request and producing party has not yet produced documents in any form.”

“Court Addresses Production of Metadata in Great Detail and Grants Production of Some but not All Data Sought,” *Electronic Discovery Law*, K&L Gates, December 2008.



# Implications for TIFF, PDF

- Most e-discovery is TIFF or PDF with a load file of metadata
  - May not be sufficient depending on needs of the case
- Native format becoming preferred
  - Avoids costs to convert to TIFF or PDF
  - Preserves substantive & embedded metadata
  - Any text errors not due to searchable text creation process

# Substantive Metadata

- Embedded in the file and moves with it
  - Created by the application
  - Track changes, comments, etc. appear in documents copied or sent to others
    - Default settings for Office 2007 programs automatically display these when open or save
    - Can be scrubbed

# System-based Metadata

- Not embedded in file - stored externally
  - Used by computer's file system to track file location
  - Stores demographics about file:
    - Name, size, creation, modification, usage
- Every active file has at least one corresponding block of system metadata
  - MS Word, PPT & Excel have 80 easily accessible substantive and system metadata fields tracked

“Beyond Data about Data: The Litigator’s Guide to Metadata, Craig Ball, [www.craigball.com/metadata.pdf](http://www.craigball.com/metadata.pdf)”



# Metadata in Litigation

- Originating path & file name
- Chain of custody
  - Name, system origin
  - Date created, modified, last accessed
    - Open doc and metadata is altered
    - Preserve early or may not be able to replicate
- Complete email header, not just To/From/Date/Subject

# Complete Email Header

- Received: from sa1.pbu.edu ([209.92.135.66]) by vms169131.mailsvcs.net
- (Sun Java(tm) System Messaging Server 6.3-7.04 (built Sep 26 2008; 32bit))
- with ESMTTP id <0KJ0005YJTYK1GD0@vms169131.mailsvcs.net> for
- **juliegable@verizon.net; Sat, 02 May 2009 09:37:33 -0500 (CDT)**
- Received: from ex1.ad.pbu.edu (www.mail.pbu.edu [::ffff:10.1.1.156])
- (TLS: TLSv1/SSLv3,168bits,DES-CBC3-SHA) by sa1.pbu.edu with esmtpp; Sat,
- 02 May 2009 10:40:53 -0400
- Received: from ex1.ad.pbu.edu ([10.1.1.156]) by EX1.ad.pbu.edu ([10.1.1.156])
- with mapi; Sat, 02 May 2009 10:40:16 -0400
- Date: Sat, 02 May 2009 10:40:16 -0400
- **From: Paula Geard <PGEARD@pbu.edu>**
- **Subject: hold on EFT**
- X-Originating-IP: [209.92.135.66]
- To: "juliegable@verizon.net" <juliegable@verizon.net>
- Message-id: <9A2A2CC0ACB39040B89FD60A03D787911519848B74@EX1.ad.pbu.edu>
- MIME-version: 1.0
- Content-type: multipart/alternative;
- boundary=\_000\_9A2A2CC0ACB39040B89FD60A03D787911519848B74EX1adpbuedu\_
- Content-language: en-US
- Accept-Language: en-US
- Thread-topic: hold on EFT
- Thread-index: AQHJyzPoEDCACbyEZUeTmCing9KnNQ==
- acceptlanguage: en-US
- X-MS-Has-Attach:
- X-MS-TNEF-Correlator:



# System Metadata Cautions

- Open or preview docs alters last access date
- Copy file changes creation date on copy
- Save file alters last modified date
- Some apps carry original author's name, even if another copies it and modifies it
- Anti-virus applications “touch” every file
- No peeking; use experts



# Relevant metadata depends on case

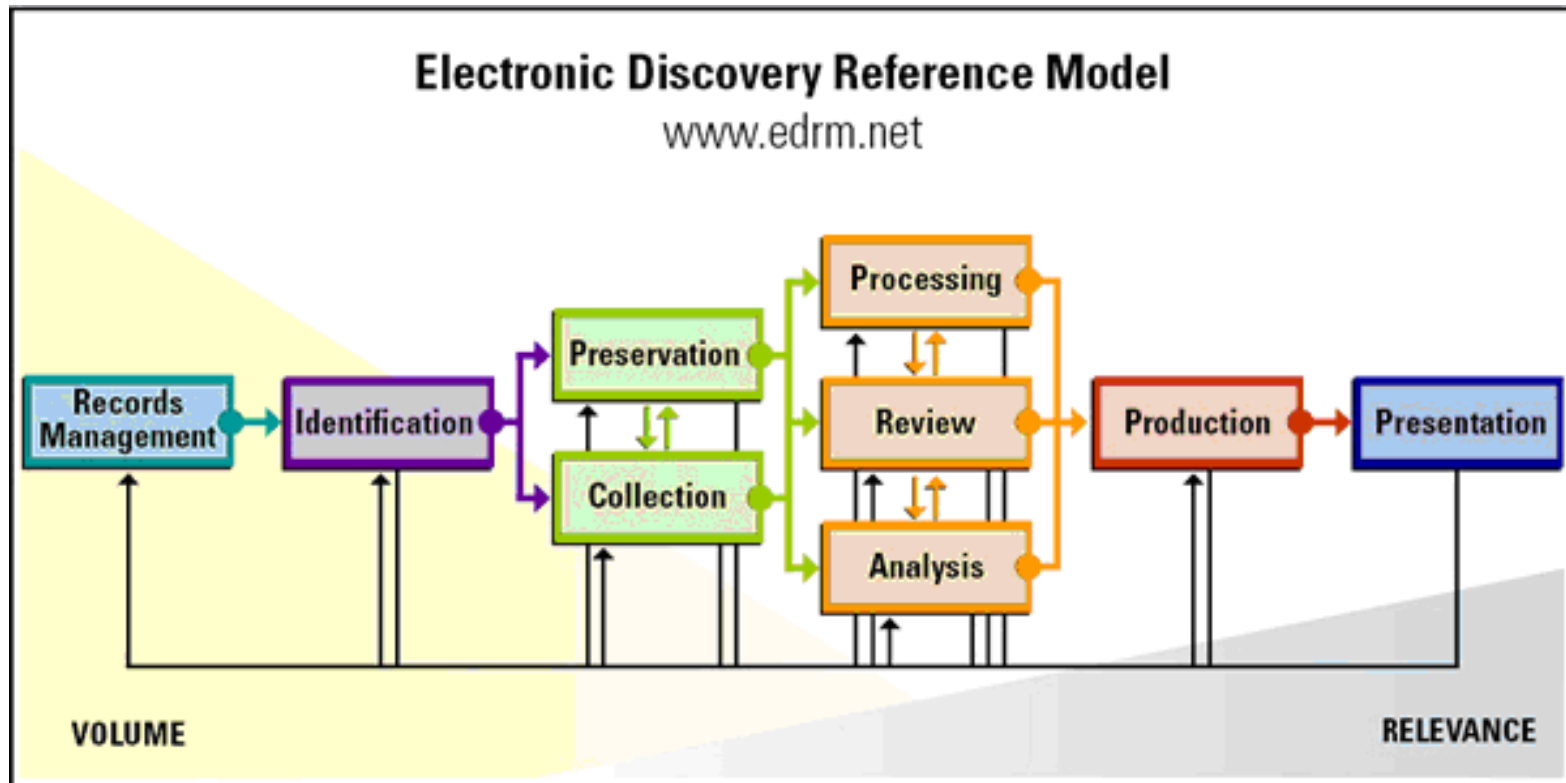
- Fraud, infringement, theft of trade secrets
  - Comparing system metadata and application metadata may show evidence of tampering
- Disputes
  - Embedded metadata may be key – formulas used to do calculations



# Authentication

- Preserve original date and system user metadata
  - Network access logs
  - Version control

# E-Discovery Reference Model



NewsWire, September 2007



# Metadata & RIM

- Metadata includes contextual, processing and use information
- Helps to identify and certify ESI scope, authenticity and integrity
- Allows organizations to better retain and organize info

# Metadata & RIM

- Absent a legal requirement to the contrary, organizations are not required to preserve metadata, but may find it useful to do so
  - Distinguish ordinary situation from litigation context (Note 51)
- Preservation of relevant metadata in litigation context
  - Migration of information from one format to another
    - Seek advice of legal counsel

# Metadata & RIM

- Preserve & migrate metadata over time
  - To survive migrations, records will need metadata that enables them to exist independently of the system used to store & retrieve them
  - Technical, descriptive and preservation metadata to show how record was created, maintained, and related to other records.

Ibid.



# Metadata & RIM

- Metadata needed to audit and track access to information
  - May be unworkable and unwise to routinely discard metadata
  - Retain sufficient metadata to ensure trustworthiness for business, fiscal, legal and historic purposes

Ibid.

# Metadata & RIM

- Organization should be aware that metadata retained in due course of business may be discoverable in its complete and original form.
  - Williams v. Sprint (2005)– Presumption in favor of producing metadata
  - Nova Measuring v. Nanometrics (2006) – Compel documents in native format with all original metadata
  - In re NYSE Securities Specialists (2006) – Court ordered all electronic docs produced in native format

Ibid.



# ISO 23081-1 Metadata for Records

## Part 1: Principles

- Guide to understanding, implementing and using metadata within ISO 15489 framework
- Addresses relevance of RM metadata in business processes
- Roles and types of metadata for RM processes
- Sets framework for managing those metadata



# Metadata at record capture

- Context of record creation
- Business context
- Agents involved
- Record content, appearance, structure and technical attributes
  - Record structure
    - Physical or technical structure
    - Logical structure, i.e., relationships between data elements comprising the record

# Metadata Management

- “Metadata about the record and metadata accruing in its management form a metadata record which must be managed”
- “It is essential to keep this metadata record at least as long as the original record exists”
- “In disposition, metadata may still be needed to account for existence”

Ibid.



# Metadata Responsibilities

- RM professionals are responsible for the reliability, authenticity, usability and integrity of metadata associated with records
- RM professionals
  - Participate in definition of metadata requirements
  - Develop related policies and strategies
  - Monitor metadata creation



# Other Metadata Areas

- Metadata for e-business
- Metadata for preservation
- Metadata for resource description
- Metadata for information retrieval
- Metadata for rights management



# Metadata Model: MoReq2

- Define metadata to allow exchange of records between ERMSs with no loss of MoReq2-related mandatory functionality
- Define such metadata sufficiently to allow the production, development, and use of a MoReq2 XML schema.

Model Requirements for the Management of Electronic Records – Update and Extension, 2008, Appendix 9: Metadata Model



# Metadata Model: MoReq2

- Describes a minimum set of metadata “elements” (aka fields) that the ERMS must be able to export, import, and process.
- Model is consistent with
  - ISO 23081 – Records management processes – Metadata for records 2
  - ISO 15489 – Records management
  - ISO 15836 – The Dublin Core metadata element set

# Metadata Architecture Research

- Description paradigm
  - Library & info science
  - Capture data about digital objects in event they are requested in user queries
- Processability and executability paradigm
  - Computer Science
  - Metadata response to user requests

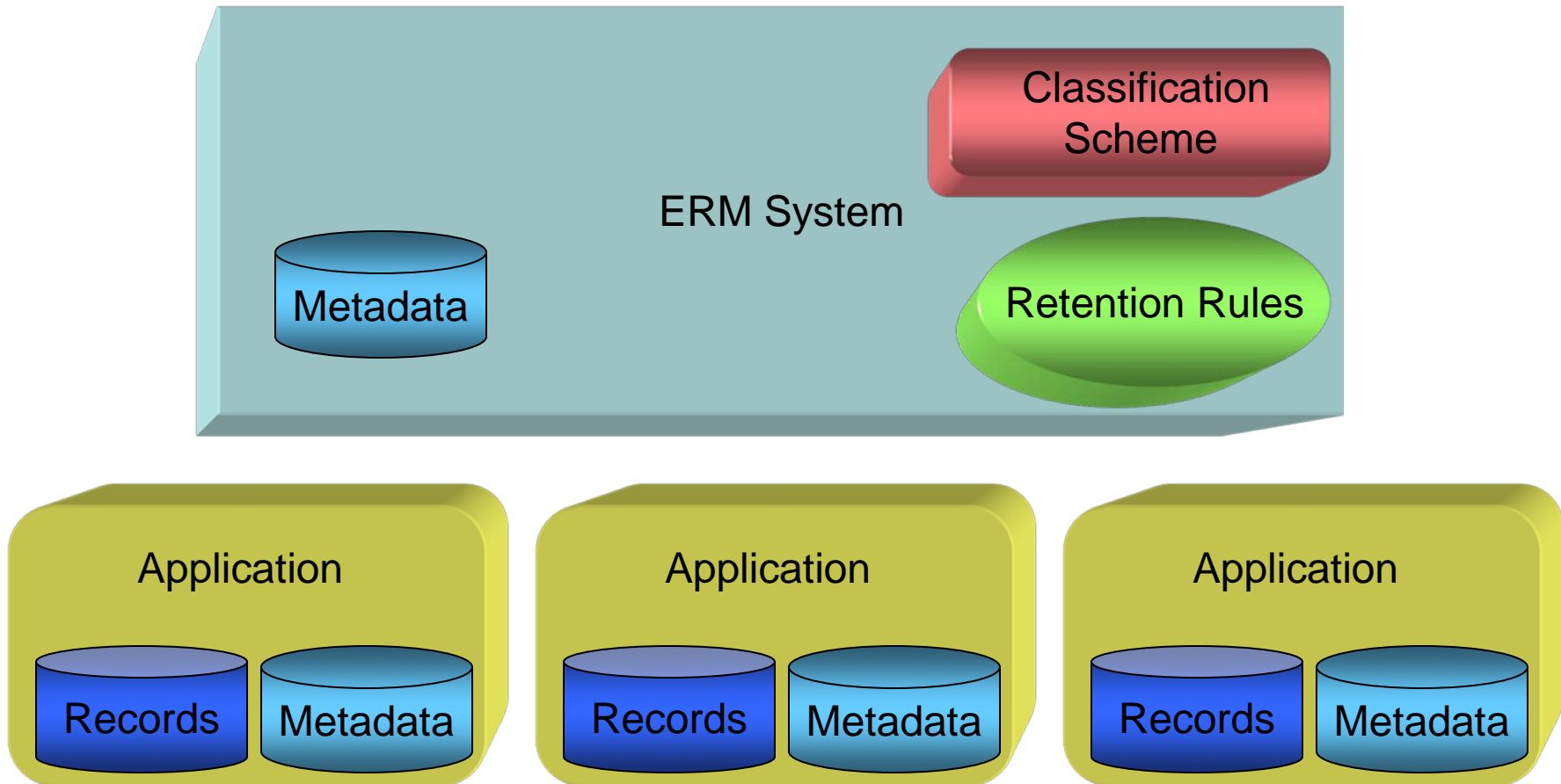
Metadata, Marcia Lei Zeng and Jian Qin, Neal-Schuman Publishers, New York, 2008



# Metadata & E-content

- Description of information assets
- Enhanced information retrieval
- Policy management & enforcement
- Ownership and authenticity
- Interoperability
- E-Discovery

# Catalog Model



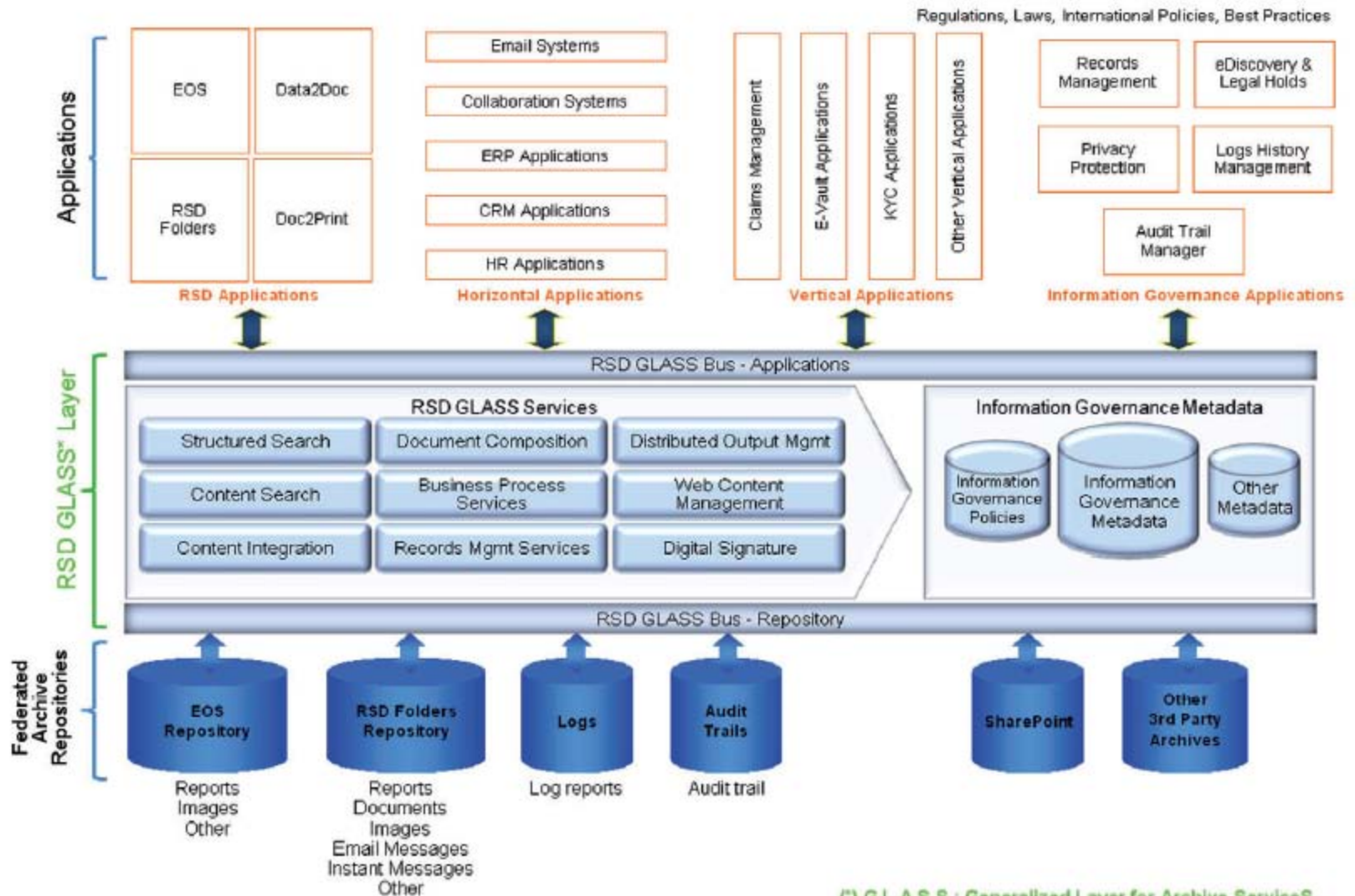


# Metadata Catalog

- Content stays in situ
- Metadata is replicated into central repository
  - Enhanced Search
  - Retention Policies
  - E-Discovery Processes

With thanks to Bassam Zarkout

# RSD Enterprise Information Governance Solution



(\*) G.L.A.S.S.: Generalized Layer for Archive Services



# Advantages

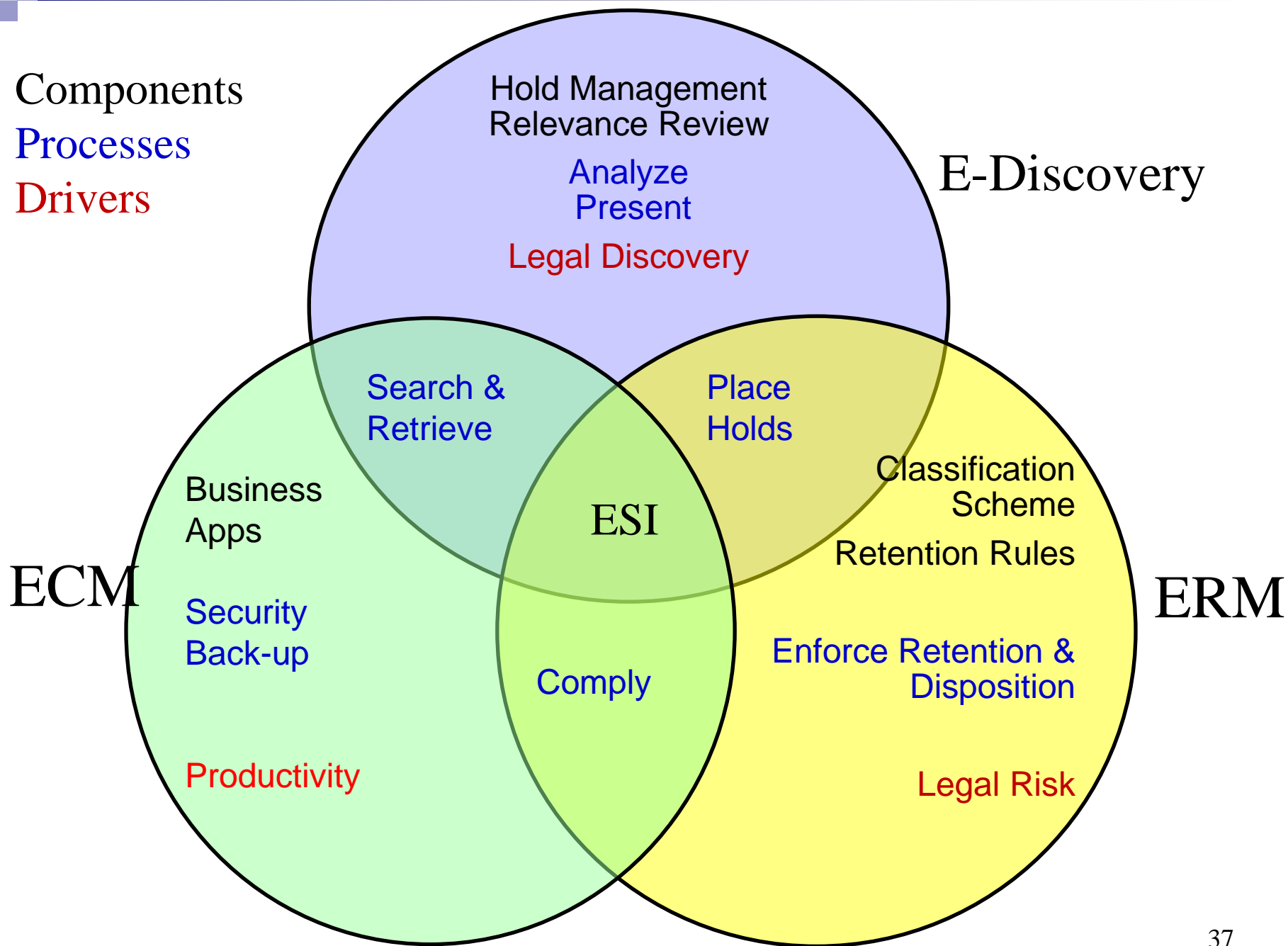
- Can handle enterprise level volumes
  - 100s to 1000s of TBs scattered across mainframes, system servers
- May offer single point of management for metadata record



# Representative Examples

- RSD GLASS – Enterprise information governance via central metadata repository
- Rational Retention – Metadata repository, services for auto-classification, e-discovery
- IBM FileNet
- Autonomy

Components  
Processes  
Drivers



	IT	Legal	Compliance	Users	RM
Metadata					
Asset Description		X		X	X
Search & retrieve	X	X	X	X	X
Info policy mgt.	X	X	X		X
Authenticity		X			X
Interoperability	X				
E-discovery	X	X			X



# Parting Thoughts

- What metadata to keep depends on many factors – including need to produce for e-discovery
- RIM responsibility to manage the metadata record
- Metadata repositories to cope with volume, diversity of electronic records

# Thank You

[juliegable@verizon.net](mailto:juliegable@verizon.net)

[www.gableconsulting.com](http://www.gableconsulting.com)

