



THE SEDONA CONFERENCE®

Commentary on Inactive Information Sources

*A Project of The Sedona Conference®
Working Group on Electronic Document
Retention & Production*

JULY 2009 - PUBLIC COMMENT VERSION



The Sedona Conference® Commentary on Inactive Information Sources:
Guidance Principles for Identifying, Classifying, Retaining and Destroying
Orphaned, Legacy and Dormant ESI

*A Project of The Sedona Conference® Working Group on Electronic Document
Retention and Production (WG1)*

Author:

The Sedona Conference®

Editor-in-Chief:

Vivian L. Polak

Senior Contributing Editors:

Jonathan Damon

Laura Owens

Deborah Juhnke

Robert Owen

Antigone Kriss

Jonathan M. Redgrave

Greg McCurdy

With input from Members of Working Group 1

Copyright © 2009 The Sedona Conference®

All Rights Reserved.

REPRINT REQUESTS:

Requests for reprints or reprint information should be directed to Richard Braman,
Executive Director of The Sedona Conference®, at tsc@sedona.net or 1-866-860-6600.

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the
members of The Sedona Conference® Working Group 1.

They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which
any of the participants belong nor do they necessarily represent official positions of The Sedona Conference®.

In addition, we thank all of our Working Group SeriesSM Sustaining and Annual Sponsors, whose support is essential to our ability to develop
Working Group Series publications. For a listing of our sponsors just click on the “Sponsors” Navigation bar on the homepage of our website.

wgsSM

Copyright © 2009,
The Sedona Conference®

Visit www.thesedonaconference.org

Preface

Welcome to another major publication in The Sedona Conference® Working Group Series: *The Sedona Conference® Commentary on Inactive Information Sources*. This effort comes from our Working Group on Electronic Document Retention and Production (WG1) and represents the collective wisdom of a diverse group of lawyers and representatives of firms providing consulting and legal services.

This effort reflects the culmination of over a year of dialogue, review and revision, including discussion at two of our WG1 meetings. In addition to the editors listed, the work product benefitted tremendously from all of the dialogue at our WG1 meetings and the comments and contributions of WG1 members on the various iterations under discussion.

On behalf of The Sedona Conference®, I want to thank the editorial team and all WG1 members whose comments contributed to this Commentary for all of their efforts to make this work product as helpful as possible.

As with all of our Working Group publications, this Commentary is first being published as a “public comment version”; after sufficient time for public comment (peer review) has passed, the editors will review the public comments and to the extent appropriate any changes in the law, and determine what edits if any might be appropriate. The Commentary will then be re-published in “final” (although in our common law system, little is ever “final.”). If you wish to submit a comment, please utilize the “public comment form” on the download page of our website.

Richard G. Braman
Executive Director
The Sedona Conference®
July, 2009

Inactive Information Guidance Principles

Inactive Information Guidance Principle 1: *Subject to any preservation obligations related to pending or reasonably anticipated litigation or government investigation, an organization should take reasonable steps to determine whether an inactive information store contains information that the organization should retain based on legal retention requirements or business needs.*

Inactive Information Guidance Principle 2: *Subject to any preservation obligations related to pending or reasonably anticipated litigation or government investigation, an organization should avoid excessive retention of inactive information by destroying it when it is no longer necessary to meet legal retention requirements or business needs.*

Inactive Information Guidance Principle 3: *An organization should take reasonable steps to determine whether an inactive information store contains information that is potentially relevant in a pending or reasonably anticipated litigation or government investigation.*

Inactive Information Guidance Principle 4: *An organization should take reasonable measures, through IT practices and user-facing policies and procedures, to reduce the ongoing accumulation of inactive information.*

Inactive Information Guidance Principle 5: *An organization should consider establishing policies and procedures for the orderly migration of data required to be retained or preserved to supported formats, systems and media to reduce the need to retain/preserve inactive information.*

Inactive Information Guidance Principle 6: *An organization should consider whether and how its policies/procedures regarding inactive information should apply to third parties in possession of the organization's inactive information.*

Inactive Information Guidance Principle 7: *An organization should consider periodically reviewing and updating any policies and procedures regarding inactive information to account for changes in laws, new forms of inactive information, and new technical capabilities or changes in business organization or requirements.*

Inactive Information Guidance Principle 8: *An organization should take reasonable steps to index/identify/organize/map corporate records (as reasonable, based on business needs) so as to minimize over-retention and disorganization.*

Table of Contents

Preface.....	i
Inactive Information Guidance Principles.....	ii
Table of Contents.....	iii
Introduction	1
I. The Problem.....	2
A. A Historical Perspective on Why the Problem Exists.....	2
1. Mainframes	2
2. Precursors to Personal Computers.....	3
3. Personal Computers	3
4. Software Development.....	3
5. Modern Electronic Communication	4
6. Increasing Volumes of Inactive Information.....	5
B. A Practical Perspective on Why the Problem Exists	5
1. Human Factors	6
2. Institutional/Organizational Factors.....	6
3. Constraints to Accessing Inactive Data	7
II. Legal Implications and Purpose	8
III. Identification, Classification, Retention and Destruction of Inactive Information Stores	9
A. Retrospectively Addressing Inactive Data—Retention Obligations	9
B. Retrospectively Addressing Inactive Data—Preservation Obligations	10
C. Prospectively Addressing Inactive Data.....	12
IV. Risk Analysis Approach to Determining the Costs/Benefits of Preserving or Disposing of Inactive Data.....	14
A. Phase 1—Pre-Investigation: Determine What You Know.....	14
B. Phase 2—Is it Reasonable to Investigate to Learn More About the Contents of the Inactive Information Store?.....	16
C. Phase 3—Learn More	17
D. Phase 4—Take Reasonable Measures to Preserve or Retain Relevant Inactive Information.....	19
V. Factors to Consider in Performing the Analysis.....	20
A. Likelihood of Relevance.....	20
1. General Questions which May Affect the Likelihood of Relevance	20
2. Specific Questions that May Increase Likelihood of Relevance	20
3. Specific Questions that May Decrease the Likelihood of Relevance	20
B. Costs of Investigating Inactive Data.....	21
1. Questions which May Point to a Lower Investigative Burden	21
2. Questions which May Point to a Higher Investigative Burden	22
C. What Steps are Reasonable to Learn More	23
1. General Questions Regarding Reasonalbe Steps to Learn More.....	23
2. Decreased Duty to Investigate	23

3. Increased duty to Investigate.....24
D. Reasonableness of Preservation or Retention Measures24
1. Decreased Duty to Take Additional Measures to Preserve or Retain.....24
2. Increased Duty to Take Additional Measures to Preserve or Retain.....25

APPENDICES

Appendix A: Decision Tree

Appendix B: The Sedona Conference® Working Group Seriessm & WGSsm Membership Program

Introduction

Since the fall of 2002, The Sedona Conference® Working Group on Best Practices for Electronic Document Retention and Production (“WG 1”) has focused on providing guidance to those faced with the challenges brought about by the exponential growth in the volume and complexity of electronically stored information found in modern enterprises. WG 1’s numerous other publications include *THE SEDONA GUIDELINES: Best Practice Guidelines & Commentary for Managing Information & Records in The Electronic Age (2nd Edition) (2007)* (“*The Sedona Guidelines*”). *The Sedona Guidelines* explore how the prevalence of electronically stored information has affected traditional concepts of records management and provide basic guidance to help organizations develop defensible processes for managing this information. *The Sedona Guidelines* provide an important background to this paper and central to it are the following notions: (1) consistent with traditional records management principles, that information should be retained as long as it has business-related value to an organization, or is required by law or regulation to be retained; (2) organizations must take steps to preserve certain information if it is “relevant”¹ in an actual or reasonably anticipated litigation event regardless of whether it constitutes a formal “record” of the organization; and (3) if information does not meet the above criteria requiring retention or preservation, then it should be destroyed.

The legitimacy of managing corporate information through records management policies that systematically destroy (as well as retain) information has been long recognized by lower courts, and has now been explicitly acknowledged by the United States Supreme Court in the case *Arthur Andersen, LLP v. United States*, 544 U.S. 696, 125 S. Ct. 2129 (2005). As the Court noted, “[d]ocument retention policies’ ... are common in business” and those policies “are created in part to keep certain information from getting into the hands of others, including the Government.” *Id.* at 2135.

This work product of WG1 builds on guidance provided by *The Sedona Guidelines* in addressing a common problem in large organizations today—information stores that persist long after their business value has expired.² We first provide some background on how such data stores are created and why they are pervasive. Then we propose a process for analyzing such information stores (on both a retrospective and a prospective basis) in the ordinary course of business and in the context of litigation.

¹ In this discussion, the terms “relevance” and “non-relevance” are used to apply to information that must be kept for ordinary course of business “retention” purposes (*i.e.*, information retained in accordance with statutory or regulatory retention requirements or internal business retention guidelines), as well as information that is subject to a preservation duty in connection with pending or reasonably anticipated litigation or investigation.

² While our primary focus is on electronically stored information (“ESI”), many organizations are also plagued with significant stores of paper records that no longer serve any useful business purpose. The guidance provided here can be used to address paper records stores as well.

I. The Problem

Many organizations have accumulated voluminous stores of ESI no longer used in the business of the organization. This information may be:

- “Orphaned” (*i.e.*, no one within the organization has knowledge of or responsibility for it);
- “Legacy” (*i.e.*, it was created by or resides on systems or programs that the organization no longer uses); or
- “Dormant” (*i.e.*, it may technically have a custodian and it may be in a format used by the organization’s current IT environment, but the information is not used or accessed).

This “inactive” information is generally, but not always, unclassified and unidentified. In some cases the organization may not have the capability to determine the contents of inactive information. It may be stored on media that is deteriorating. It may be difficult to access because it is stored on obsolete hardware or was created using obsolete software and the organization may no longer have the technical equipment or personnel needed to access or make use of it. And, in many cases this inactive information has simply been forgotten.

Despite the fact that the organization no longer actively uses the inactive information and may never use it again, the organization continues to store it indefinitely—often due to inertia, a lack of awareness that the information stores exist or a perceived low return on investment for the time and expense of separating and discarding inactive information —increasing data storage costs and raising the potential that the information may become subject to preservation and production obligations in the event of litigation or government investigation.

Addressing this problem has both a backward-looking aspect (to address historic or accumulated data) and a forward-looking aspect (to address the creation of corporate policies and procedures designed to reduce future accumulation of inactive information). The methods of addressing the problem will vary based on a particular organization’s circumstances and will change depending on when the analysis is conducted—*i.e.*, when the analysis is undertaken in the ordinary course of business, the decision-making process will be different from the process undertaken when an organization anticipates or is involved in litigation or a government investigation.

A. A Historical Perspective on Why the Problem Exists

In the last six decades the IT industry has been one of the fastest growing and most dynamic sectors in economic history, literally changing the world. Technology has become increasingly efficient in creating, processing, transmitting and storing electronic information, dramatically increasing the number of people using computer resources and the volume of electronically stored information. The effects of increased efficiency in processors, data storage, and software have been further magnified by the proliferation of institutional networks, the Internet and the Web.

1. Mainframes

The first computers, developed in the World War II era, were generally owned by only the very largest institutions. These (huge) mainframe computers were expensive to purchase (often costing \$1 million or more) and maintain. They were accessed mainly through “dumb terminals”—user interfaces with little or no data processing or storage capacity. Consequently, very few people worked with computers.

Mainframes were offered by the likes of IBM, Bull, Burroughs, Data General, Fujitsu, ICL, Nixdorf and Sperry-Rand. Hardware and software products of these various manufacturers were incompatible with each other. Data generated and stored on one such system could not be stored or used on a different system without significant difficulties. This resulted in increased volumes of data that were difficult to manage or access when organizations migrated from one provider's technology to another, or even upgraded to new technology provided by the same company.

2. Precursors to Personal Computers

Since those early days, innovation and competition in the IT sector have led to ever smaller, cheaper and more powerful computers. In the 1970s "minicomputers" were developed to assist with engineering and scientific tasks. The 1980s brought a number of "microcomputer" offerings—precursors to today's Personal Computers ("PCs"). But again, computers from different vendors were incompatible, with each running its own operating systems. Operating systems and the processors on which they run must be designed for each other, or they will not work. Additionally, software applications written for one brand of computer would not run on any other brand, making it difficult to share information. Because of the lack of interoperability, upgrades to or replacements of these technologies meant the creation of new stores of inactive information. However, the total volume of inactive information remained small because relatively few of these computers were sold. That would soon change. By the mid-1980s, IBM introduced its first true PC, setting in motion the wheels that led to the data accumulation explosion seen today.

3. Personal Computers

IBM's first PC was offered with a choice of three operating systems, but over time Microsoft's MS-DOS became the most popular operating system on IBM PCs. Other PC manufacturers soon imitated IBM's technology, including chips designed to work with MS-DOS and similar operating systems. Eventually, Windows was introduced and its successive versions became the operating system and processor standard for IBM and IBM-compatible PCs ("Windows PCs").

Apple introduced its Macintosh computer as an integrated product (*i.e.*, Apple provided the hardware and software) and from 1984-1999 Apple maintained its proprietary closed operating system standard generally referred to as "Mac OS." Mac OS was incompatible with IBM and IBM-compatible PCs. Beginning in 1999 and over the next several years, Apple began transitioning all its new Macintosh computers to "OS X," a new operating system based in part on the UNIX standard to improve compatibility with other technologies. In 2006, Apple switched to Intel-based microprocessor architecture for its new PCs seeking to improve compatibility between Macintosh and IBM-compatible PCs by permitting Microsoft Windows (and other Intel-based operating systems) to be installed and run on Macintosh computers. But this did not resolve all of the compatibility issues between these computers with fundamentally different operating systems.

4. Software Development

Software has undergone even greater development, experimentation and proliferation over the last six decades than hardware.

In the 1940s computer users wrote machine code manually. During the 1950s-1960s rudimentary tools were developed to increase productivity and improve quality. During the 1970s more sophisticated tools with collaborative features were developed (*e.g.*, Unix). Small business software started to be developed to serve the minicomputers that were on the rise. During the 1980s as PCs became popular, consumer software was developed. The introduction of object-oriented programming, the dramatic increases in computer memory

capacity, and the significant drop in PC prices during the 1990s meant the demand for software grew. The internet increased the general availability of software. This led to the development of platforms to make software development easier (*e.g.*, Java) through the early part of the 21st Century, with offshore outsourcing lowering the cost of application development and adding to the overall amount of software available for home and business.

In his book *In the Beginning: Recollections of Software Pioneers*, (1998) IEEE Computer Society Press, author Robert L. Glass divides software engineering into three time periods: the Pioneering Era, the Stabilizing Era and the Micro Era.³ According to Glass, during the Pioneering Era (1955-1965) as new computers came out every couple of years, old computers were rendered obsolete and all software programs had to be re-written to run on new computers. This was the era of “machine rooms,” punch cards, job-queues and I-O (in-out) windows. Software was written for a specific type of computer hardware, with different disciplines requiring not only different software, but different machines. Eventually languages like FORTRAN and COBAL were developed to ease the frequent need to translate previously developed software to newly developed computers. During this era, no one was selling pre-packaged software. Computer companies generally provided software free with their computers and users shared it freely.

The Stabilizing Era (1965-1980) was heralded by the IBM 360, which combined business and scientific applications into one computer, offering both binary and decimal arithmetic. As the largest software project of that time, it put an end to the constant parade of newer-cheaper computers every year or so. While programmers still developed code in machine rooms and had to learn a whole new language to communicate with the operating system (*i.e.*, job control language or JCL), they were able to begin writing new software instead of retrofitting old software for new computers. In this era, the demand for computer programmers exceeded the available supply, and by the late 1960s some schools began to offer computing as an academic subject. The concept of “Structured Programming” developed along with standards organizations. Some software vendors began to emerge, but most customized applications continued to be developed in-house. The concept of software engineering was eventually developed and in 1968 the NATO Conference on Software Engineering was held—greatly influencing the field.

The Micro Era (1980-1998) was characterized by dropping prices, and software engineering became a known and respected career choice for programmers with desktop machines. As the difficult to use JCL was replaced by the graphical user interface (“GUI”), user interaction with electronic devices, including computers, became easier.

Today new technologies are introduced at an incredibly rapid pace. As of 2008, there are more than 2,500 documented programming languages⁴ and literally hundreds of software vendors have come into being with thousands of offerings. While the software industry has made great efforts to provide compatibility between older and newer products, as technology marches ahead in software and hardware, incompatibilities inevitably arise.

5. Modern Electronic Communication

Although the Department of Defense and academic institutions had been developing integrated computer networks for years, even by the 1980s the distribution of electronic documents was slow and cumbersome with most people printing and distributing documents on paper or exchanging floppy disks or diskettes.

³See <http://www.cs.colorado.edu/~kena/classes/5828/s99/comments/srinivasan/01-29-1999.html> (last visited February 15, 2008).

⁴See “The Language List” maintained by Bill Kinnersley, <http://people.ku.edu/~nkinners/LangList/Extras/langlist.htm> accessed February 15, 2008. See also the poster entitled “History of Programming Languages”, http://www.oreilly.com/pub/a/oreilly/news/languageposter_0504.html (last visited February 15, 2008)

The 1990s saw the development of communications protocols that carried data across the wires and increasing development of client-server networks.⁵ What followed were wide-area networks, corporate intranets, and the Internet—which collectively served to connect many previously unconnected PCs.

Today large institutional IT networks have become extremely complex—often connecting thousands of servers and tens of thousands of clients.⁶ The composition of each of these networks is also constantly changing as the hardware and software are upgraded and adapted to evolving business needs. As with PCs, server software and hardware products rapidly age until they become obsolete, with manufacturers eventually terminating support or updates. The data then becomes increasingly difficult to access and use unless migrated to new systems.

Following on the heels of technologies designed to more easily create and store electronic data came the development of technologies to more easily communicate such data (*i.e.*, email and instant messaging). This has exponentially increased the amount of information that is recorded and saved (even temporarily).

6. Increasing Volumes of Inactive Information

Inactive data is often generated when an organization moves from one technology platform—hardware, software or communications—to another and retains archival versions of data from the previous system, while discarding the hardware it would require to access this information. Even incremental upgrades during routine replacement of PCs or servers can create caches of inactive information if the hardware is retained and does not have its hard drive wiped before it is placed in storage. One highly visible source of inactive information is backup tapes that are kept well past the time of their usefulness from a disaster recovery perspective and that if kept long enough, may even have been created using technology no longer in use at an organization.

The result of all this technological innovation: ever-increasing stores of ESI, including inactive ESI. The Manual for Complex Litigation (4th) § 11.446 is illustrative (albeit somewhat dated):

The sheer volume of such data, when compared with conventional paper documentation, can be staggering. A floppy disk, with 1.44 megabytes, is the equivalent of 720 typewritten pages of plain text. A CD-ROM, with 650 megabytes, can hold up to 325,000 typewritten pages. One gigabyte (1000 megabytes) is the equivalent of 500,000 typewritten pages. Large corporate computer networks create backup data measured in terabytes, or 1,000,000 megabytes: each terabyte represents the equivalent of 500 billion typewritten pages of plain text.

B. A Practical Perspective on Why the Problem Exists

ESI accumulates in organizations for many different reasons, making the identification of all potential sources of even active data exceedingly difficult. Repositories of ESI can be as physically small as a flash drive, floppy disk or CD-ROM and may be mixed with hardcopy materials stored in file cabinets or off-site storage boxes. Inactive information is often stored haphazardly or stricken from inventory lists. Physical storage devices lose their

⁵ The general concept of a client-server network is simple, but its many implementations can be extremely complex. “Clients,” in IT jargon, are commonly PCs including desktops and laptops, but can also include non-PC devices that individuals use—such as Blackberries, Pocket PCs, and other PDAs (personal digital assistants). These clients can all be used to create and store data. Thanks to networks and the Internet, they can also be used to send and receive data. “Servers,” on the other hand, typically serve, that is, process and deliver data to client computers connected by a network. There are many different kinds of servers in terms of their hardware and the operating systems and applications software that run on them. These diverse servers are designed and deployed to perform many different functions and are used for tasks such as serving up web pages, emails, and other file and data types.

⁶ Different businesses or organizations (*e.g.*, banks, airlines, government agencies, retailers, software companies, or government agencies) have very different communications and data needs and thus use very different networks running many different types of software and hardware on both their clients and servers. Mainframe computers that are still used in many large IT systems also act as servers in networks. Database servers specialize in storing vast amounts of data in many different forms.

functionality and the data stored on them loses its business value. As individuals leave an organization or move about within an organization, fewer remaining people remember the location and purpose of data repositories. Thus, technical, institutional and human factors all lead to the accumulation of inactive information.

1. Human Factors

A common source of inactive information is the working files of individuals in an organization. These accumulations of inactive information may be known only to the individuals who created them, and generally are not tracked by centralized information management protocols and systems.

At present most organizations do not employ formal document management systems solutions. Rather, employees create and maintain electronic documents on their workstation hard drives or network server locations (both shared and individually assigned). While individuals often develop their own naming and storage conventions, it is not uncommon to discover that the organization has no formal system for either. As an employee moves on to other projects or even a new organization, the context and value of materials created in the past may become lost and the materials become inactive information.

Working files are also often stored on removable media easily carried from one location to another. Whether out of an abundance of caution or a failure to appreciate the problems caused by indiscriminate and rampant retention of data, many individuals keep these storage media as part of their personal files (as opposed to official files of the company) rather than discarding them or making sure that the applicable contents are properly retained in company-approved storage.

The amount of this accumulated inactive information, the time pressures of the modern business world and human nature often conspire against current employees taking the time to go back and delete inactive information. If the employee leaves the organization, or is re-assigned to a different location, this data becomes orphaned, and the review of such information by others to determine its usefulness is often a low priority. It continues to exist merely by virtue of neglect, and may eventually take on the characteristics of legacy data or orphaned data should the passage of time render it difficult to access.

2. Institutional/Organizational Factors

Inactive information also accumulates when organizations reorganize or merge with other organizations. Particularly in the case of a corporate merger, identifying and consolidating the electronic information that the combined organization will use going-forward routinely creates large volumes of older data that are no longer useful to the organization. This older data may be consciously archived for long-term storage, or it may simply accumulate at the fringes of corporate activity.

Litigation events are another source of inactive information. Many times, failing to properly consider alternative approaches, an organization will preserve backup tapes or other disaster recovery storage media in the face of pending or “reasonably anticipated” litigation. This data initially may be accessible using tools and expertise within the organization, but often exists as a large mass of undifferentiated files and data. Moreover, many organizations have exposure to overlapping litigation matters, in which materials collected for one matter cannot be released because they may contain material possibly relevant in another, newer matter. Such situations create a “cascading litigation hold” situation in which increasingly large amounts of inactive information continues to be

retained solely to satisfy the obligation to preserve potentially relevant data. Over time, this inactive information can become not only unmanageably voluminous, but may also be stored on obsolete media no longer in use by the organization.⁷

3. Constraints to Accessing Inactive Data

Inactive information may be difficult and costly to restore or reconstruct for a variety of reasons, including degradation of the media on which it is stored⁸ or technological changes that have taken place in the time since the information was created. Even if the required hardware or software is available and information archives have been converted to searchable form, it is often difficult to identify the content of such archives without opening the data files themselves.⁹ It can therefore be difficult and costly to properly classify documents and make retention decisions. Also, many of these archives may have been created by individuals who are no longer with the organization, making first-hand knowledge of the archives' content simply unavailable.¹⁰

Not surprisingly, reviewing inactive information may lack urgent business justification. If a business is fully functional without accessing this data, it is difficult to quantify and demonstrate the return on investment that proactive review of these materials would generate. Indeed, proactive review of inactive information is most commonly justified by comparing the cost of such review to the costs of collection, review, and production (as well as the risk of potential liability) that would be incurred if these materials become potentially relevant in a legal dispute. While these costs are real, many organizations faced with tight budgets find it difficult to justify the expenditure in advance of an actual litigation event.

⁷ See *The Sedona Principles, Second Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Discovery* 24, available at <http://thesedonaconference.org>, (discussing the costs of retaining and managing superfluous information).

⁸ For tapes, floppy disks, and other media using magnetic particles to store information, the electrical charge of the oxide particles storing data gradually weakens over time, making it impossible to access this data. A common and particularly challenging example is backup archives, which generally comprise discrete, unclassified and perhaps unrelated files. When magnetic tapes are used as the storage medium, mechanical and magnetic deterioration of the backup media is a significant concern. Similarly, the data stratum of CD-ROMs and DVD-ROMs degrades over time, sometimes making the archived files unreadable in as little as a year after their creation. Thus, even carefully preserved data with no prior history of data errors may prove to be unreadable when queried.

⁹ Older software typically did not create or track much of the metadata that is associated with modern digital files (such as when a given file was accessed and under which computer login ID). Therefore, even data stored in searchable form may lack the external metadata that makes it possible to categorize files without directly viewing their contents.

¹⁰ This problem is exacerbated by simple human nature—the fact that few users of removable media take the time to properly label each piece of media with information as to when it was last used and its contents.

II. Legal Implications and Purpose

We take it as a given that human examination of all data to assess classification, retention and destruction is simply not possible in light of the volume and complexity of electronic data today. Thus, we provide here some basic guidance to help organizations, courts and regulators understand the problem and process for assessing the reasonableness of actions that have been or should be taken to address the issues.

We are hopeful that this document will serve as a catalyst for a “tipping point” where more order and rationality is applied to the generation and storage of data to meet retention and preservation obligations in good faith.¹¹ It is also our desire that hindsight review of retention and destruction activities for their reasonableness will be undertaken with an understanding of the principles and considerations set forth in this paper.

¹¹ Additional guidance in the litigation context is provided in *The Sedona Conference® Commentary on Legal Holds* (August 2007 Public Comment Version).

III. Identification, Classification, Retention and Destruction of Inactive Information Stores

Different considerations and priorities will apply to an organization's determination of how to manage inactive information when it is under a duty to preserve evidence for litigation or government investigation than when the determination is made in the ordinary course of business.¹² In order to distinguish between preservation of evidence in connection with litigation or government investigation, as opposed to retention of information in the ordinary course of business, the following nomenclature is used herein:

- Retention—determination in the normal course of business of how to manage information considering statutory, regulatory, and/or business requirements.
- Preservation—compliance with a duty to preserve information related to a pending or reasonably anticipated litigation, government investigation or audit.

A. Retrospectively Addressing Inactive Data—Retention Obligations

Inactive Information Guidance Principle 1: Subject to any preservation obligations related to pending or reasonably anticipated litigation or government investigation, an organization should take reasonable steps to determine whether an inactive information store contains information that the organization should retain based on legal retention requirements or business needs.

Inactive Information Guidance Principle 2: Subject to any preservation obligations related to pending or reasonably anticipated litigation or government investigation, an organization should avoid excessive retention of inactive information by destroying it when it is no longer necessary to meet legal retention requirements or business needs.

Organizations should make reasonable and good faith efforts to identify existing sources of inactive information that are subject to legal retention requirements, both within the organization and in the possession of third parties. Generally, organizations will not need to physically review the entire contents of an inactive information store in order to make reasonable, good faith and defensible decisions as to whether it contains information that the organization may be required to retain.

What is considered “reasonable,” and the specific steps taken during the process for identifying whether an inactive information store contains information that the organization should retain, will vary depending on the circumstances. One reasonable approach is to perform a risk analysis. This analysis is focused on using available or reasonably obtainable information to assess the likelihood that the inactive information store might contain data that the organization should retain. Specifically,

- If an organization can identify, with little to no effort, monetary cost, or other resource expenditures, that the inactive information store is unlikely to contain information that the organization should retain, the organization may destroy the inactive information store without further investigation.
- If a cursory review of the inactive information store cannot confirm an absence of information that should be retained, it may be necessary for the organization to take additional measures to attempt

¹² As noted below, the overall structure of the analysis and several specific elements of this guidance will apply in both the ordinary course of business context and in connection with the preservation of evidence that may be relevant in litigation or an investigation.

to learn more about the inactive information content before it can determine whether the inactive information store contains information that the organization is required to retain.

- If an organization determines that additional investigation of the inactive information contents is warranted, it only needs to take such specific information-gathering measures as are reasonable under the circumstances to determine if the information stored must be retained.

When evaluating the reasonableness of the investigation into the contents of an inactive information store, one should take into consideration the costs and burdens associated with the steps needed to learn more about the information store. If the information store is relatively inaccessible (*i.e.*, while it is technologically feasible to access the contents of the information store, doing so would be extremely burdensome and costly), it may be unreasonable to take additional steps to access the information unless the likelihood that it contains information that must be retained is extremely high. On the other hand, if the burden and cost associated with accessing the information in order to learn more is low, this should be factored into the decision-making. Even if the information store is easily accessible, it may be very difficult or costly to learn about the contents (*e.g.*, if the inactive information is extremely voluminous, and unsearchable, unsortable, unindexed and not organized in any way). In these cases, an organization may be limited to methods of learning more about the information store that do not involve accessing its contents (*e.g.*, institutional knowledge), if any such methods are available.

If the organization conducts a good faith, reasonable investigation and does not find that the inactive information store contains information that the organization has a legal requirement (based on statutory or regulatory obligations) or a business need to retain and is not subject to preservation obligations arising out of litigation or government investigation, the organization is under no duty to continue to retain the inactive information store and may destroy it. Organizations should take into consideration privacy and information security concerns, including legal requirements for secure destruction of certain types of information, when destroying inactive information. Prior to final disposition, organizations should document the key steps taken to arrive at the determination that the organization did not have a duty to retain the inactive information, including ensuring that actions are taken consistent with applicable records retention schedules in effect.

If an organization determines that the inactive information store is reasonably likely to contain information that the organization has a legal requirement to retain, the organization should ensure that reasonable steps, as well as any specific measures mandated by applicable statutory or regulatory requirements, are taken to maintain the inactive information in usable or readable form.¹³

B. Retrospectively Addressing Inactive Data—Preservation Obligations

Inactive Information Guidance Principle 3: An organization should take reasonable steps to determine whether an inactive information store contains information that is potentially relevant in a pending or reasonably anticipated litigation or government investigation.

In the litigation and government investigation context, organizations generally have the obligation to preserve, and possibly to review and produce, relevant information in their possession, custody or control. Thus, organizations must make reasonable and good faith efforts to identify existing sources of inactive information, both within the organization and in the possession of third parties over whom they have control, that may contain information potentially relevant in a pending or reasonably anticipated litigation or government

¹³ Statutory or regulatory requirements regarding the accessibility of information may mandate that data be moved from deteriorating media to ensure that such data is maintained in a usable/readable format for the duration of its retention period. *See, e.g.*, 36 C.F.R. 1234.30(g)(4) (federal regulations require agencies to re-copy permanent data residing on magnetic tapes media onto new media every ten years).

investigation.¹⁴ Again, generally organizations will not need to physically review the entire contents of an inactive information store in order to make defensible and reasonable decisions in this regard. Whether or not the organization's efforts in this regard are reasonable or in good faith will vary depending on the circumstances.

When an organization is under a duty to preserve potentially relevant information there may be additional considerations that should be addressed in the risk analysis, above and beyond those described above in connection with normal retention obligations. These include:

- There may be a duty to preserve multiple copies of each file or document (*e.g.*, in different formats, with handwritten notes, showing receipt by different people, etc.).
- The risk of destroying relevant inactive information includes potentially significant spoliation sanctions.
- Agreements with opposing parties and/or preservation orders issued by courts or regulators may mandate specific investigative and/or preservation steps with respect to a particular inactive information store, superseding any independent risk analysis or assessment of reasonable steps that the organization would otherwise conduct.

An organization may destroy inactive information, if, based on a good faith and reasonable investigation, the organization has a reasonable belief that an inactive information store does not contain potentially relevant or responsive information and if the organization does not violate a preservation agreement or order by doing so.

Absent agreement with opposing parties or a preservation order from a court or regulator specifying the preservation measures to be taken, the duty to preserve extends to taking reasonable steps to preserve potentially relevant or responsive inactive information. Stated differently, this means that organizations are not under an obligation to take "heroic efforts" to preserve an inactive information store, even if it contains or is reasonably likely to contain, potentially relevant or responsive information. While the general standard is the same as in the retention context, the additional considerations in the preservation context may alter the outcome.

A reasonableness/proportionality analysis should be conducted to determine whether it would be reasonable under the circumstances to take steps to preserve a specific inactive information store that contains, or is reasonably likely to contain, potentially relevant or responsive information that should be preserved. If so, a determination should be made as to what steps would be "reasonable." Some of the factors that may be considered include:

- Relative accessibility of the inactive information;
- Likelihood that the information contained in the inactive information store is unique; and
- The risks associated with failing to preserve the inactive information.

If potentially relevant and unique inactive information exists in a medium that is in the process of deteriorating such that the information will become irretrievable or unreadable while subject to a preservation obligation, the organization must weigh the added costs of taking affirmative steps to ensure that the data remains retrievable and readable against the degree of potential relevance or responsiveness and the uniqueness of the information. The organization should confer with opposing parties and/or the court or regulator early in the process if the organization (based on good faith, reasonable investigation) determines that the inactive information store likely

¹⁴ Of course, any retention obligations applicable to inactive information continue to apply in the preservation context and are not diminished or superseded by the additional preservation obligations that arise in connection with pending or reasonably anticipated litigation, government investigation, audit or subpoena.

(a) contains potentially relevant information; (b) that is unique; (c) and not otherwise available from another source; (d) but that requires an “undue burden” to preserve.¹⁵ Cost shifting may be appropriate if the organization is required to undertake an “undue burden” or make “heroic efforts” to preserve an inactive information store.

Under the Federal Rules of Civil Procedure as amended in 2006 to specifically discuss discovery of ESI, litigating parties may be required to identify all data sources by category and type (including inactive information) that may contain potentially relevant or responsive information that they will rely on to support their claim/defense or that would be likely to lead to the discovery of admissible evidence. This includes those sources that the party considers to be “not reasonably accessible because of undue burden or cost.”¹⁶ Accessibility may affect whether inactive information will be relevant in litigation. If, in the ordinary course of business, an organization routinely alters and/or deletes data in an inactive information store (e.g., for retention purposes), this may weigh against any claim that the information store is not reasonably accessible. Such routine alteration or deletion may affect the extent of the duty to preserve an inactive information store.

C. Prospectively Addressing Inactive Data

Inactive Information Guidance Principle 4: An organization should take reasonable measures, through IT practices and employee policies and procedures, to reduce the ongoing accumulation of inactive information.¹⁷

Inactive Information Guidance Principle 5: An organization should consider establishing policies and procedures for the orderly migration of data required to be retained or preserved to supported formats, systems and media to reduce the need to retain/preserve inactive information.¹⁸

Inactive Information Guidance Principle 6: An organization should consider whether and how its policies and procedures regarding inactive information should apply to third parties in possession of the organization’s inactive information.

Inactive Information Guidance Principle 7: An organization should consider periodically reviewing and updating any policies and procedures regarding inactive information to account for changes in laws, new forms of inactive information, and new technical capabilities or business requirements.

Inactive Information Guidance Principle 8: An organization should take reasonable steps to index/identify/organize/map corporate records (as reasonable, based on business needs) so as to minimize over-retention and disorganization.

There are numerous direct and indirect costs and risks associated with the unbridled accumulation and retention of inactive information.¹⁹ Organizations should take a good faith, reasonable approach to retention,

¹⁵ For example, if costly affirmative measures, such as rehabilitating deteriorating media, will have to be taken to ensure potentially relevant or responsive information in the inactive information store remains retrievable and readable.

¹⁶ Fed. R. Civ. P. 26(b)(2)(B) (2006). See also *Sedona Principles Second Edition 18* (“Importantly, the Rules do not require the identification of all inaccessible sources of electronically stored information, but only those that the producing party believes in good faith may contain relevant, non-duplicative information.”).

¹⁷ Organizations should consider addressing issues related to inactive information in their record retention programs, and establishing user-facing policies and procedures providing guidance on steps that end-users, managers, and IT personnel can take to reduce the accumulation and over-retention of inactive data. For example: (a) User-facing guidance regarding when to dispose of data and how to ensure data is really deleted; (b) Human resources policies on how to deal with data upon employee termination or transfer; (c) Policies for IT personnel on recycling backup tapes when they are no longer needed for disaster recovery; and (d) Procedures for dealing with inactive information acquired through extraordinary circumstances, including mergers and acquisitions.

¹⁸ See also *Sedona Guidelines 84* (“As new applications are developed or acquired within organizations, the records management requirements relative to those applications should be anticipated and planned as part of the system development or software and/or hardware selection. Digital preservation requires routine efforts to migrate records to overcome software and technological obsolescence and from deteriorating media.”).

¹⁹ See *Sedona Guidelines 24* (“Retaining superfluous electronic information has associated direct and indirect costs and burdens that go well beyond the cost of additional electronic storage.”).

management and disposition of inactive information stores both on a prospective basis (to avoid the future accumulation inactive information) and on a retrospective basis (to manage existing inactive information).

Organizations should consider taking a multi-disciplinary approach to dealing with inactive information, taking into account legal requirements, business and IT interests, and the organization's interest in avoiding retention of superfluous information.²⁰ No single standard or model for retention, preservation, management or disposition of existing or future inactive information can meet every organization's needs, or even a single organization's needs under every set of circumstances.

²⁰ See *Sedona Guidelines* 24 ("At the heart of a reasonable information and records management approach is the concept of the "lifecycle" of information based on its inherent value. In essence, this means that information and records should be retained only so long as they have value as defined by business needs or legal requirements.").

IV. Risk Analysis Approach to Determining the Cost/Benefits of Preserving or Disposing of Inactive Data

The “reasonable” steps that organizations may take to identify whether an inactive information store contains information that the organization should retain or preserve will vary depending on the circumstances. The analysis focuses largely on costs of reviewing and analyzing the contents of an inactive information store versus the likelihood that those contents may be “relevant.” The risk analysis must be broad enough to be useful in determining whether inactive information is subject to retention requirements as well as in identifying inactive information that may fall within the scope of a preservation duty. However, as discussed above, there may be differences in the way the risk analysis is applied, or in the factors considered, in the preservation context as opposed to the retention context.

The approach outlined below (and reflected in the accompanying decision tree) uses a cost/benefit risk analysis to determine what constitutes a reasonable process for identifying whether inactive information may be disposed of or whether it must be retained or preserved.²¹ The approach is set forth in four separate phases so that organizations may determine, as they learn more about a particular inactive information store, whether it is reasonable to proceed to the next phase or whether the costs of attempting to continue to the next phase outweigh the potential benefits.

A. Phase 1—Pre-Investigation: Determine What You Know

Without expending significant resources, the organization should determine what is known about the likelihood that the inactive information store will contain information that the organization should retain or preserve. If this cursory review results in a reasonable determination, under the circumstances, that the inactive information does not contain information that must be retained or preserved, the organization may destroy the inactive information.

The first phase involves determining:

- Information immediately available about the inactive information store, or that can be obtained using preliminary measures, such as by physical inspection of the inactive information store;
- Information available through institutional knowledge; and/or
- Information otherwise ascertainable with little to no effort, monetary cost, or other resource expenditures.

In some cases, an inactive information store will have a label or may be associated with a legible or viewable index or inventory that will allow a fast determination of relevance or non-relevance of the inactive information. There may also be enough institutional knowledge regarding the content and purpose of an inactive information store to provide sufficient information for making an immediate relevance determination. There may be other immediate indicators of non-relevance (*e.g.*, it was created as a temporary backup copy of information available from more active sources; it pertains to a defunct project that is pertinent to no litigation, is not subject to ongoing legal or business retention or preservation requirements, and has no applicability to the current business of the organization). (*See* Illustrations 1.A. and 1.B.)

²¹ *See Sedona Principles Second Edition, supra note 7, Principle 8 and comments 45-48* (addressing “not reasonably accessible” information and balancing of burden and costs versus potential relevance).

Illustration: 1.A

Company A has email records for its employees going back ten years. The company is actively involved in litigation with relevant documents going back from 2 to 8 years. Two years ago, Company A converted to a new email system. During the migration, the prior email was saved in its old software system on three servers. These old servers are examples of legacy information stores. The company's IT department has determined that one of the servers only has emails that are 9 and 10 years old, and the records officer of the company has determined that the company has no statutory or regulatory obligations to retain documents that are older than 8 years. Because the information on that server is not relevant in any litigation and not subject to any retention obligations, Company A does not have to proceed to Phase 2 (Reasonableness of Investigation) with respect to the one server with 9-10 year old emails. Instead, it may document the reasons why the inactive information is no longer needed and may dispose of that server.

Illustration 1.B:

Company B has a big warehouse with shelves and tubs full of old backup tapes. The backup tapes are approximately 10 to 20 years old and the company no longer has the hardware, software, or personnel needed to access them. The company has no legal requirements, business need or any anticipated or pending litigation or government investigation that relate to documents from the 1980s. To the extent the company can identify tapes from the 1980s, these can be disposed of without going to Phase 2 (Reasonableness of Investigation). The company does, however, have a few litigations in which documents from the early 1990s may be relevant. The tapes from 1990 to 1995 have labels identifying which servers they backed up, but the labels do not reveal what data was on the various servers. The company has no index or description of which servers held what data in that time period, and no current company personnel have detailed memory of those servers or their contents. For the tapes from the 1990 to 1995 time period, preservation obligations indicate that Company B should proceed to Phase 2 (Reasonableness of Investigation). If there are no legal or statutory retention requirements applicable to the tapes that pre-date 1990, Company B may document the process used to determine that it had no obligations to retain such tapes and then may dispose of them.

In many cases the inactive information store does not present sufficient information on its face and there is not sufficient institutional knowledge to enable a relevance determination without some investigation into the inactive information store. In such cases, it will be necessary to determine whether any additional steps should be taken to learn more about the likely content of the inactive information store (and potential relevance of such content). This second, more common, situation is addressed in Illustrations 1.C and 1.D, below.

Illustration 1.C:

Company C finds a backup tape that contains no label and determines that the hardware that was used for creating, reading and restoring that backup tape was used for several years in the 1980s but has not been available or functional since 1992. The institutional knowledge regarding that particular type of tape is limited, except that it is known that that type of tape was used for multiple purposes during the 1980s. Some of the uses included storing a backup of a server, a large database, and an email system. In this situation, Company C knows virtually nothing about this tape and may discard the tape, unless it has a requirement to keep data from prior to 1992 (due to legal requirements or because of a pending or reasonably anticipated litigation or government investigation), in which case it should proceed to Phase 2 (Reasonableness of Investigation).

Illustration 1.D:

Company D finds an old server that is known to contain a prior email system no longer in use. The company knows of 100 users who were on that email system, 5 of whom have been identified as potentially possessing information relevant in a pending regulatory investigation. The time period that the email system was in use is known. However, the "relevant" time period for the investigation has not been identified by the regulatory agency or agreed upon by the company and the agency. Although Company D has gathered

some information about the inactive information store, it is not enough to determine relevance without further knowledge regarding the timeframe at issue in the regulatory investigation. Company D should proceed to Phase 2 (Reasonableness of Investigation). However, depending on the circumstances (e.g., the costs of proceeding to Phase 2 (Reasonableness of Investigation), the cost of continuing to preserve the server), Company D may wish to request or await clarification from the regulator regarding the timeframe at issue in the investigation.

B. Phase 2—Is it Reasonable to Investigate to Learn More About the Contents of the Inactive Information Store? (Reasonableness of Investigation)

If the Phase 1 (Determine What You Know) review of inactive information cannot confirm an absence of information that should be retained or preserved, it may in some cases be reasonable for the organization to take additional measures to learn more about the content of the inactive information store. In other cases, the cost of learning more about the inactive information contents will outweigh the likelihood that the inactive information contains relevant information. Phase 2 (Reasonableness of Investigation) provides a mechanism for weighing these factors to determine whether it is reasonable to conduct further investigation of the inactive information contents. Phase 2 (Reasonableness of Investigation) also provides a mechanism for determining that the likelihood of finding information that should be retained or preserved is so low that the risk of inadvertent destruction of potentially relevant information is outweighed by the benefit of destroying the inactive information store.

In cases where the content of the inactive information store is not immediately discernible, organizations may undertake additional steps to learn more about the inactive information store in order to make a determination of relevance. For example, organizations may wish to perform search functions, review metadata, create indexes and reports, or take other additional steps to learn about the inactive information store (e.g., information categories, types, time periods, custodians, file-related software programs, uses, and current availability). The process for determining what types of information should be collected for this inactive information store and how to collect such information is discussed in Phase 3 (Learn More). In some cases, however, the costs of gathering additional information that would enable the organization to make a relevance determination with respect to a particular inactive information store will be unreasonably high, and such costs may outweigh the benefits in certain circumstances, such as where the likelihood that the inactive information store contains relevant information is determined to be low in Phase 1 (Determine What You Know). In these cases, it may be unreasonable to proceed to Phase 3 (Learn More).

Phase 2 (Reasonableness of Investigation) is intended to help organizations determine whether it is reasonable to attempt to investigate or analyze the inactive information store in order to gather information not readily discernible in Phase 1 (Determine What You Know) regarding the potential relevance of the contents of the inactive information store. Going through Phase 2 (Reasonableness of Investigation) will help a company determine whether or not it is reasonable to proceed to Phase 3 (Learn More). Such an investigation, however, will be case-specific and will depend on a number of factors. These factors may provide additional information regarding the likelihood that the inactive information store will contain relevant information and may inform the decision as to whether the organization should proceed to Phase 3 (Learn More).

The reasonableness analysis will center on the answers to two key questions:

- **What is the likelihood of relevance based on known facts** (e.g., does the inactive information store pertain to a key player in a pending litigation; does it relate to an ongoing or core business of the organization)?

- **What is the difficulty/cost of obtaining additional information** (*e.g.*, is the inactive information store organized, indexed or part of an “electronic landfill”; is the inactive information store on legacy systems that are no longer operational or commercially available; has the employee responsible for the inactive information left the organization)?

The relative weights of these issues can be used to determine whether the organization should take additional reasonable steps to learn more about the inactive information store (*i.e.*, proceed to Phase 3) or whether such additional steps would constitute an unreasonable, undue burden and the inactive information can be disposed of without further investigation. To conduct the risk analysis, the organization should weigh the likelihood of relevance of the contents of the inactive information store against the factors relating to the costs/difficulty of investigating the inactive information store. To assist in this risk analysis, some of the factors to consider are included in non-exhaustive lists attached at the end of this document (*see* Section V, Phase 2 Factor List).

If, after undertaking the analysis described above, a determination is made that the investigative burden outweighs the potential relevance (*i.e.*, that further investigation would be unreasonably burdensome), then it would be reasonable for the organization to dispose of the inactive information (after carefully documenting the analysis in Phases 1 and 2 prior to any such disposal).²² If a determination is made that the burden of investigation is reasonable in light of the likelihood that the inactive information store contains potentially relevant information and that additional facts about the contents of the inactive information store can be obtained with relative ease, the organization should proceed to Phase 3 (Learn More).

C. Phase 3—Learn More

If the Phase 2 (Reasonableness of Investigation) analysis indicates that further investigation of the inactive information contents is warranted, Phase 3 (Learn More) provides a mechanism for determining which investigative steps would be reasonable under the circumstances before the costs of further investigation outweigh the likelihood that the inactive information will contain relevant information. Phase 3 (Learn More) also provides a mechanism for assessing the additional information such that a reasonable determination can be made that the risk of inadvertent destruction of relevant information is outweighed by the benefit of destroying the inactive information store.

If an organization determines that some additional investigative and information-gathering measures should be taken as a result of the outcome of its analysis in Phase 2 (Reasonableness of Investigation), it should determine what would constitute “reasonable steps” to learn more about the inactive information store. Such steps could include:

- Identifying the information on the inactive information store by category, type, author, custodian, or time period;
- Interviewing key custodians;
- Interviewing persons with historical knowledge of the organization’s systems;
- Referring to any pre-existing content index;
- Creating an index;

²² Disposal may not be an option in the litigation or investigation context, where finding potential relevance may trigger a duty to preserve (although not necessarily to search or produce, if the costs of search and production are determined to be unduly burdensome) and/or where agreements with opposing parties or preservation orders of a court or regulator may obviate a party’s need/ability to conduct its own risk analysis.

- Conducting key word searches;
- Sampling by accessing individual files for further investigation;
- Accessing and reviewing metadata; and/or
- Accessing and sorting the contents of the inactive information store.

The determination of precisely which additional investigative steps should be taken will be case-specific, and will vary depending on the type of inactive information as well as on a number of other factors. Some of these factors are included in a non-exhaustive list that can be found at the end of this document as Phase 3 Factor List (“What Steps are Reasonable?”).

1. Taking Investigative Steps

In many cases, the organization should begin its investigation with less costly steps that are likely to yield partial results rather than immediately taking more costly steps (such as restoring a previous server environment to reboot old backup tapes). In determining the initial step to take, organizations should weigh the cost of the investigative step against the likelihood that it will yield useful information. The concept of reasonable partial review measures for purposes of evaluating relevance has been applied by several courts; courts have allowed, and in some cases ordered, sampling, pre-approved key word searches and similar measures.²³ Steps that are so costly that their burden outweighs their expected benefits would likely not be considered reasonable. For example, for data that is on degraded media, searching the data may not be an option.

The determination of which steps to take during Phase 3 (Learn More) (and how far to go in the investigation) will likely be informed by the outcome of the analysis in Phase 2 (Reasonableness of Investigation). For example, the likelihood that the inactive information is relevant, as determined by Phase 2 (Reasonableness of Investigation), may determine the size of the “reasonable” sample chosen for Phase 3 (Learn More).

- a. If the results of the analysis in Phase 2 (Reasonableness of Investigation) weigh heavily in favor of proceeding to Phase 3 (Learn More) due to a high likelihood of relevance, the organization should continue taking investigative measures until it reaches the point where the burden of taking further measures would outweigh the likely value of such measures.
- b. If the results of the analysis in Phase 2 (Reasonableness of Investigation) and the factors relevant to Phase 3 (Learn More) do not unequivocally indicate that there is a duty to preserve or retain the inactive information (perhaps because there was a low likelihood of relevance and a low investigative burden) the organization should consider other low-cost investigative steps.
- c. Each time an organization learns something during the investigative process, the additional information should be incorporated into the organization’s assessment of its investigative duty. Each new piece of information could show:
 - i. Enough increased relevance that the organization determines the inactive information is relevant, in which case the organization should proceed to Phase 4 (Reasonable Measures to Preserve);

²³ See, e.g., *McPeck v. Ashcroft*, 202 F.R.D. 31, 33 (2001) (“There is certainly no controlling authority for the proposition that restoring all backup tapes is necessary in every case.”).

- ii. Some added relevance but not enough to proceed to Phase 4 (Reasonable Measures to Preserve) (in which case, the increased likelihood of relevance could be enough to increase the investigative duty, obligating the organization to take further actions under Phase 3 (Learn More)); or
- iii. Decreased relevance (which may decrease the extent of the investigative duty and allow the organization to cease its investigation at Phase 3 (Learn More) and conclude it must either retain/preserve the information or that it may destroy it).

If these steps produce nothing new, it may be reasonable to refrain from employing additional, higher cost steps for which the investigative burden would outweigh the expected benefits. If, however, the organization takes low-burden steps and discovers information indicating the presence of potentially relevant information in the inactive information store, it should factor this new information into its investigative duty analysis and perhaps alter its view of what efforts it should undertake in its investigation.

D. Phase 4—Take Reasonable Measures to Preserve or Retain Relevant Inactive Information (Reasonable Measures to Preserve)

If the Phase 3 (Learn More) analysis leads to a determination that the inactive information may contain relevant information sufficient to warrant retention/preservation efforts, Phase 4 (Reasonable Measures to Preserve) addresses how an organization may determine what additional steps, if any, are “reasonable” to maintain the inactive information in a way that the contents will continue to be available (even if not “reasonably accessible”). In the retention context, regulators will often dictate measures that must be taken so that information subject to regulatory retention requirements is maintained under certain conditions to ensure its availability, accessibility or legibility. In the preservation context, absent a preservation order from a court or regulator or an agreement between the parties that covers this issue, the standard for “reasonable steps” to preserve information (including information that might not be “reasonably accessible” or that is deteriorating) is based on a general common law duty of a party to pending or reasonably anticipated litigation or government investigation to preserve potential evidence.

Generally, the analysis of what measures should be taken to preserve or retain relevant inactive information is based on weighing a number of factors (*e.g.*, the uniqueness of the information on the inactive information store, the importance of information), a non-exhaustive list of which is included as Factor List 4 (“Reasonableness of Preservation or Retention Measures”) at the end of this document. This Factor List 4 has a high degree of overlap with considerations for whether information is “reasonably accessible.”²⁴ However, the determination of what constitutes a reasonable plan for preservation/retention of inactive information may differ, depending on whether the analysis is done in the context of litigation or a government investigation, as opposed to analysis in the ordinary course of business retention based on legal requirements and business needs. In the litigation/government investigation context, parties will often have to share information about data stores, including inactive information stores, with opposing counsel, an investigating agency and/or the court during some stage of the inactive information relevance determination process. In these situations, the analysis process will be fundamentally the same as that discussed here, except that the risks may be higher and the final decision as to what approach will be taken may be the product of agreement with opposing counsel (or by order of a court or regulator).

²⁴ Readers may also wish to review *The Sedona Conference® Commentary on Preserving and Managing Information that Is Not Reasonably Accessible*.

V. Factors to Consider in Performing the Analysis

A. Phase 2 Factor List (A): Likelihood of Relevance

The following non-exhaustive list of factors can help a company determine the likelihood that an inactive information store contains relevant information. (*See* discussion of Phase 2 (Reasonableness of Investigation) for application of these factors.)

1. General Questions which May Affect the Likelihood of Relevance

- a. Do the files relate to current or ongoing business of the organization?
- b. Does the inactive information store relate to a core business of the organization?
- c. To what time period does the inactive information pertain?
- d. Why was the inactive information retained or preserved?
- e. Was it retained due to an affirmative decision to retain or preserve?
- f. Is the inactive information currently subject to another litigation hold that is still active?
- g. Is there any relationship between the focus of that litigation hold and the reason why the inactive information is currently subject to this investigation?

2. Specific Questions that May Increase the Likelihood of Relevance

This non-exhaustive list of factors, if answered in the affirmative, may indicate an increased likelihood of relevance:

- a. Has the inactive information been recently and/or regularly accessed in the course of the organization's business?
- b. Has the inactive information been retained pursuant to a document retention schedule?
- c. Is the inactive information content in a format likely to contain mostly user-created information (such as word processing documents, presentations, emails, and spreadsheets)?
- d. Was the inactive information created by, or does it relate to, a business activity that is the focus of an ongoing or anticipated litigation or government investigation?
- e. Was it created or modified during the timeframe of interest in a pending or reasonably anticipated litigation or government investigation?

3. Specific Questions that May Decrease the Likelihood of Relevance

If answered in the affirmative, the following non-exhaustive list of factors may indicate a decreased likelihood of relevance:

- a. Is the medium unlikely to contain user-created data (*e.g.*, is it a backup of an application server or a server containing primarily systems and software?)
- b. Is the inactive information replicated elsewhere?
- c. Does the inactive information essentially amount to an inactive copy of an active information store (*e.g.*, the inactive information constitutes a static pre-migration copy of information otherwise available)?
- d. Is the inactive information a backup that contains duplicate data available from other existing and accessible sources?
- e. Is the inactive information old enough that its contents have exceeded all of the retention periods that the company assigns based on legal/business requirements (*e.g.*, if the company's longest retention period is 50 years and the inactive information is more than 50 years old) or the applicable litigation(s)/government investigation(s) in question? If so, the inactive information may be subject to a date cut-off determined by the applicable relevance issue – *e.g.*, if the relevance issue is a litigation that pertains to business conducted by the organization in the 1990s and the inactive information contains files from 1975.
- f. Date cut-offs may also be used for inactive information pertaining to departments, teams, projects or entities that ceased to do business for the organization prior to the time period at issue, or for inactive information belonging to employees terminated prior to the relevant time period.
 - For example, if the inactive information relates to work of a former employee no longer with the organization, how much time has passed since the individual was employed with the organization? If the employee was terminated prior to the relevant time period, a date cut-off may be appropriate.
 - If the inactive information relates to departments, teams, projects or entities (*e.g.*, sold or merged companies) that no longer do business for the organization and ceased doing business prior to the relevant time period, a date cut-off may be appropriate.

B. Phase 2 Factor List (B): Costs of Investigating Inactive Data

This non-exhaustive list of factors relates to the “investigative burden,” *i.e.*, the relative difficulty or ease of learning additional information regarding the inactive information.

1. Questions which May Point to a Lower Investigative Burden

If answered in the affirmative, the following non-exhaustive list of factors may indicate a lower investigation burden:

- a. Can the information on the inactive information store be identified by category, type, author, custodian, time period or other useful characteristics without further effort?

- b. Are the tapes, boxes, bins, shelves, machines, etc., containing the inactive information labeled in any way?
- c. Is there a pre-existing content index that provides details sufficient to enable an informed determination of relevance?
- d. If not, can an index be created at a reasonable cost?
- e. If an index exists or can be created, are the costs of reviewing the index reasonable?
- f. Is there some useful organizational structure to the inactive information?
- g. Is it possible to retrieve information about the contents of the inactive information without actually accessing it? (Note: accessibility costs should be strongly considered in the risk analysis because access costs may play an important role in determining the overall investigative burden.)
- h. Is the inactive information store currently searchable and/or able to be sorted?
- i. Is it possible to sample the inactive information store by accessing individual files for further investigation?
- j. Is there metadata available regarding the content of the inactive information store, and if so, is it possible to access and review the metadata?
- k. Are the tools (hardware equipment and software) needed to access the inactive information and conduct further investigation on its contents available to the organization?
- l. If the tools needed to access and investigate the inactive information are not available to the organization, are they readily available through other sources?

2. Questions which May Point to a Higher Investigative Burden

If answered in the affirmative, the following non-exhaustive list of factors may indicate a higher investigative burden:

- a. Is the inactive information “organized” or simply an “electronic landfill”? (Note: if the inactive information is an “electronic landfill,” this may greatly increase the investigative burden.)
- b. Are there multiple inactive information stores in one physical location where an analysis of one inactive information store requires further analysis of all (*i.e.*, is it difficult or impossible to separate individual inactive information stores for more focused investigations)?
- c. Is the inactive information on media or hardware susceptible to deterioration or decay?
- d. Is it necessary to retrieve and review the inactive information in order to determine its contents? If so:

- Can the inactive information be retrieved or reviewed at all (*e.g.*, deteriorating media problem)?
- What is the cost of retrieving and reviewing inactive information? For example, can one simply insert a disc into a drive or will it be necessary to recreate a prior server environment in order to restore backup tapes?²⁵

C. Phase 3 Factor List: What Steps are Reasonable to Learn More

To determine how far an organization should go in its investigation into the contents of the inactive information store and what would be “reasonable” under the circumstances (*i.e.*, at what point do the costs of investigating further outweigh the benefits), several factors should be considered.²⁶

1. General Questions Regarding Reasonable Steps to Learn More:

- a. How recently has the inactive information been accessed or used?
- b. Has it been accessed, used or retrieved since it became inactive information and if so when and how many times?
- c. Is the custodian or creator of the information still employed with the organization?
- d. Are the software programs associated with the inactive information files still available at the organization?
- e. Are the software programs associated with inactive information files still commercially available at a reasonable cost?
- f. Is the inactive information source a backup tape that is used strictly for disaster recovery, and is not accessed in the ordinary course of business (*e.g.*, when executives want to find deleted emails)? Can it be associated with the files of “key players”?
- g. If the software programs associated with the inactive information are not available to the organization, are there other programs or methods that are available that can be used to access the information? How reasonable are the costs of these alternatives?

2. Decreased Duty to Investigate

If present, the following non-exhaustive list of factors may favor decreased scope of, and resource expenditure on, additional investigation into relevance of the inactive information contents:

- a. Informed predictions as to the importance and usefulness of inactive information indicate a low likelihood that it will be important or useful for the relevant purposes (*e.g.*, central to litigation issues, important for core business activities);
- b. If investigated for determining the duty to preserve, the likely probative value of the contents of the inactive information store for the issues at stake in the pending or anticipated litigation or government investigation is low;

²⁵ See Fed. R. Civ. P. 26(b)(2)(B) (2006 Advisory Committee Note).

²⁶ See also, *Sedona Principles Second Edition*, supra note 7, Principle 5. This flowchart and the steps discussed here are meant to help determine what actions are “reasonable” within the meaning of this Principle.

- c. A high quantity of relevant information is known to be available from other and more easily accessed and analyzed sources; and
- d. It is unlikely that the relevant information on the inactive information store is unique.

3. Increased Duty to Investigate

If present, the following non-exhaustive list of factors may favor increased scope of, and resource expenditure on, additional investigation of the inactive information contents:

- a. A high likelihood of relevance;
- b. A determination that the information contained in the inactive information store cannot be obtained from other, more easily accessed and analyzed sources;
- c. A low burden of identifying relevant information available from the inactive information store (and, in some cases, of preserving and retrieving that subset of information on the inactive information store that is relevant); and
- d. A high likelihood that any relevant information on the inactive information store is unique.

D. Phase 4 Factor List: Reasonableness of Preservation or Retention Measures

To determine the measures that would be “reasonable” for an organization to take in order to meet its obligations for retention or preservation of inactive information in a given set of circumstances, several factors should be considered.

1. Decreased Duty to Take Additional Measures to Preserve or Retain

If present, the factors on the following non-exhaustive list may indicate a decreased duty to take additional measures to retain and/or preserve relevant inactive information:

- a. A high quantity of relevant information is known to be available from other, more easily accessed and analyzed sources;
- b. The information is inactive information and likely to significantly degrade or deteriorate over time;
- c. The inactive information is likely to become more costly or burdensome to recover over time;
- d. It is unlikely that the relevant information in the inactive information source is unique (*i.e.*, it likely exists elsewhere);
- e. Informed predictions as to the importance and usefulness (*e.g.*, is it central to litigation issues or retention for statutory or regulatory needs) of the inactive information-derived information indicate a low likelihood that it will be important or useful for the relevant purposes; and

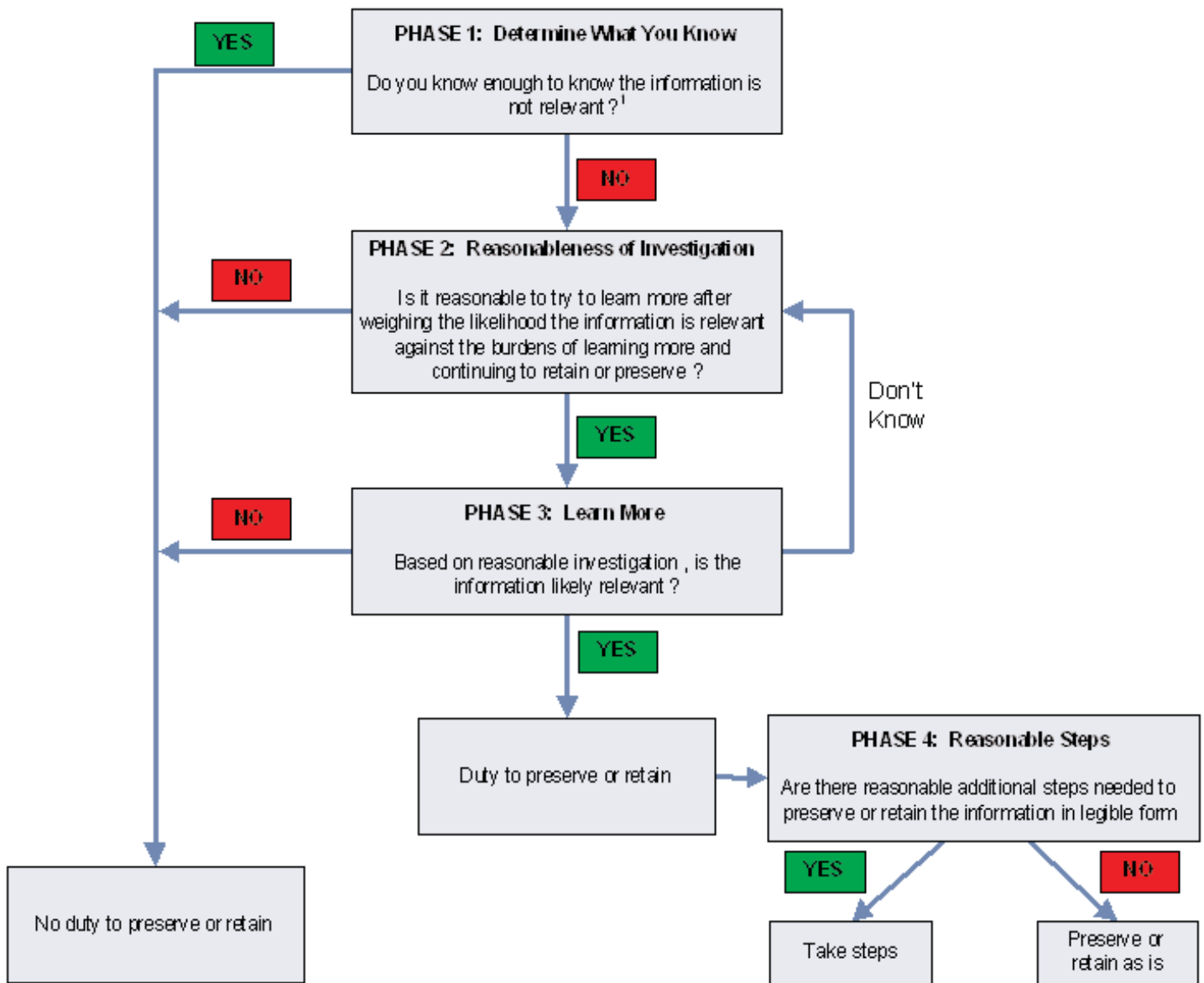
- f. If investigated in the preservation context, the information in the inactive information source is unlikely to have much probative value for the issues at stake.

2. Increased Duty to Take Additional Measures to Preserve or Retain

If present, and as informed by the determinations in Phase 3, the factors on the following non-exhaustive list may indicate an increased duty to take additional measures to retain and/or preserve relevant inactive information:

- a. The inactive information is highly accessible;
- b. The information in the inactive information source cannot be obtained from other, more easily accessed and analyzed sources; and
- c. Relevant information can be easily filtered from irrelevant information in the inactive information source, and if necessary, separating, retaining, preserving and producing only that pocket of relevant information is feasible without undue burden or cost.

Appendix A: Decision Tree



1 In this diagram, the term "relevant" is used to apply to information that must be kept for ordinary course of business "retention" purposes (i.e., information retained in accordance with statutory or regulatory retention requirements or internal business retention guidelines), as well as information that is subject to a preservation duty in connection with pending or reasonably anticipated litigation or investigation.

Appendix B: The Sedona Conference® Working Group Series & WGSSM Membership Program

“**DIALOGUE
DESIGNED
TO MOVE
THE LAW
FORWARD
IN A
REASONED
AND JUST
WAY**”

The Sedona Conference® Working Group Series (“WGSSM”) represents the evolution of The Sedona Conference® from a forum for advanced dialogue to an open think-tank confronting some of the most challenging issues faced by our legal system today.

The WGSSM begins with the same high caliber of participants as our regular season conferences. The total, active group, however, is limited to 30-35 instead of 60. Further, in lieu of finished papers being posted on the website in advance of the Conference, thought pieces and other ideas are exchanged ahead of time, and the Working Group meeting becomes the opportunity to create a set of recommendations, guidelines or other position piece designed to be of immediate benefit to the bench and bar, and to move the law forward in a reasoned and just way. Working Group output, when complete, is then put through a peer review process, including where possible critique at one of our regular season conferences, hopefully resulting in authoritative, meaningful and balanced final papers for publication and distribution.

The first Working Group was convened in October 2002, and was dedicated to the development of guidelines for electronic document retention and production. The impact of its first (draft) publication—The Sedona Principles; Best Practices Recommendations and Principles Addressing Electronic Document Production (March 2003 version)—was immediate and substantial. The Principles was cited in the Judicial Conference of the United State Advisory Committee on Civil Rules Discovery Subcommittee Report on Electronic Discovery less than a month after the publication of the “public comment” draft, and was cited in a seminal e-discovery decision of the Federal District Court in New York less than a month after that. As noted in the June 2003 issue of Pike & Fischer’s *Digital Discovery and E-Evidence*, “The Principles...influence is already becoming evident.”

The WGSSM Membership Program was established to provide a vehicle to allow any interested jurist, attorney, academic or consultant to participate in Working Group activities. Membership provides access to advance drafts of Working Group output with the opportunity for early input, and to a Bulletin Board where reference materials are posted and current news and other matters of interest can be discussed. Members may also indicate their willingness to volunteer for special Project Team assignments And a Member’s Roster is included in Working Group publications. Please go to our website —www.thesedonaconference.org—for an application to join the Working Group Series.SM

We currently have active Working Groups in the areas of 1) electronic document retention and production; 2) protective orders, confidentiality, and public access; 3) the role of economics in antitrust; 4) the intersection of the patent and antitrust laws; (5) Markman hearings and claim construction; (6) international e-information disclosure and management issues; and (7) e-discovery in Canadian civil litigation. See the “Working Group Series” area of our website www.thesedonaconference.org for further details on our Working Group Series and the Membership Program.