

# REPORT

## ADMISSIBILITY, ELECTRONIC EVIDENCE & INFORMATION MANAGEMENT COMPLIANCE:

### AN EVALUATION OF EMC CENTERA

KAHN CONSULTING, INC.

RANDOLPH A. KAHN, ESQ.  
BARCLAY T. BLAIR

---

**TABLE OF CONTENTS**

---

<b>TABLE OF CONTENTS .....</b>	<b>2</b>
<b>I. EXECUTIVE SUMMARY.....</b>	<b>3</b>
SUMMARY OF EVALUATION .....	3
EVALUATION OVERVIEW .....	3
<b>II. INTRODUCTION.....</b>	<b>4</b>
THE INFORMATION MANAGEMENT CHALLENGE .....	4
LEGAL ADMISSIBILITY AND EVIDENTIARY STRENGTH.....	5
<b>III. ABOUT CENTERA.....</b>	<b>6</b>
CENTERA FEATURES.....	6
<i>Content Addressing</i> .....	6
<i>Redundant Array of Independent Nodes (RAIN)</i> .....	7
<i>Content Protection Mirroring (CPM) and Content Protection Parity (CPP)</i> .....	7
<i>Data Regeneration: "Self Healing"</i> .....	7
<b>IV. CENTERA CAPABILITIES.....</b>	<b>8</b>
1. LONG-TERM CONTENT INTEGRITY .....	8
2. CONTENT COMPLETENESS AND AUTHENTICITY .....	9
3. CONTENT ACCESSIBILITY .....	10
4. INFORMATION SECURITY .....	11
5. RESISTANCE TO DELETION AND OVERWRITING.....	11
6. RECORDS RETENTION.....	12
7. BUSINESS CONTINUITY AND DISASTER RECOVERY .....	12
8. RECORDING AND STORAGE PROCESS INTEGRITY AND AUTHENTICITY .....	13
9. RECORDS DISPOSITION .....	13
<b>VI. ABOUT KAHN CONSULTING.....</b>	<b>15</b>

---

## I. EXECUTIVE SUMMARY

---

### SUMMARY OF EVALUATION

It is the opinion of Kahn Consulting, Inc. that EMC's Centera™ Compliance Edition provides a compelling platform for the trustworthy storage of electronic records and other digital information required for legal and regulatory purposes. By protecting the integrity, reliability, accessibility, and accuracy of information throughout its lifespan, Centera can play an important role in helping organizations retain and manage information in manner that will better ensure its admissibility and promote its evidentiary strength. By protecting content from alteration and unauthorized deletion; by verifying the accuracy of information during the recording process; by providing long-term content accessibility; by supporting disaster recovery and information security needs; and by supporting records retention and disposition functionality, Centera promotes the authenticity and trustworthiness of electronic records and digital evidence. Furthermore, Centera's capabilities can assist organizations to comply with certain information management and compliance requirements contained in laws such as the Sarbanes-Oxley Act of 2002.

### EVALUATION OVERVIEW

Kahn Consulting, Inc. ("KCI") was engaged by EMC Corporation ("EMC") to evaluate the company's Centera Compliance Edition storage platform ("Centera"). The primary purpose of this Evaluation is to assess the product's utility as a platform for the retention of electronic records and other digital information required for legal and regulatory purposes. Rather than focusing only on specific laws or regulations, in conducting this Evaluation KCI has assessed Centera functionality against criteria derived from broad legal and regulatory requirements for admissibility, electronic evidence, and records management. Retaining and managing digital information in manner that will satisfy the courts and regulators depends on a proper program of technology, people, and technical and procedural controls. This Evaluation assesses the value that Centera may bring to such a program.<sup>1</sup>

KAHN CONSULTING, INC.

157 LEONARD WOOD NORTH • HIGHLAND PARK IL • 60035

PHONE: 847.266.0722 • FAX: 847.266.0734 • E-MAIL: INFO@KAHNCONSULTINGINC.COM

---

## II. INTRODUCTION

---

### THE INFORMATION MANAGEMENT CHALLENGE

*“The world’s total production of information amounts to about 250 megabytes for each man, woman, and child on earth. It is clear that we are all drowning in a sea of information. The challenge is to learn to swim in that sea, rather than drown in it. Better understanding and better tools are desperately needed if we are to take full advantage of the ever-increasing supply of information. . .”*

*UC Berkeley study<sup>2</sup>*

Today’s organizations rely on information technology in practically every facet of their business.<sup>3</sup> As a result, more and more of the information that organizations use to operate their businesses and to comply with laws and regulations exists in electronic form. A recent study showed that 93% of all corporate documents are created electronically, and that printed documents comprise only .003% of the total amount of information worldwide.<sup>4</sup> At the same time, the sheer volume of digital information created and stored by organizations continues to grow. For example, the number of email messages sent per day is expected to grow from 31 billion in 2002 to 60 billion by 2006.<sup>5</sup>

At the same time, organizations are increasingly motivated to ensure that electronic records and other evidence of their business activities are captured, retained, and managed in a manner that will satisfy the courts and regulators, as well as their own business needs. There are a variety of factors driving this, including:

- **Increasing reliance on electronic information.** More and more of the information that organizations require to operate their businesses and to comply with laws and regulations today is created and managed in electronic form. In fact, according to the courts, “[c]omputers have become so commonplace that most court battles now involve discovery of some type of computer-stored information.”<sup>6</sup> Much of the information that once would have existed in paper form now exists solely in electronic form, and may never be reduced to printed form.
- **Heightened scrutiny.** The high-profile business failures and stories of corporate malfeasance that have filled the trade and mainstream media over the past few years have contributed to an environment of increased legal and regulatory scrutiny of information management practices. The result is that regulators, courts, and shareholders today demand accountability in the way that organizations capture, retain, and manage their digital information.
- **Laws and regulations.** The law today enables organizations to rely on electronic records for more purposes than ever before. However, while the law may have evolved to accommodate digital business processes, legal requirements to ensure that electronic records and evidence are secure, authentic, complete, and reliable have not. As such, organizations need tools and controls that can help them manage the information that they capture and retain for business and legal purposes in a trustworthy manner.

KAHN CONSULTING, INC.

157 LEONARD WOOD NORTH • HIGHLAND PARK IL • 60035

PHONE: 847.266.0722 • FAX: 847.266.0734 • E-MAIL: INFO@KAHNCONSULTINGINC.COM

## LEGAL ADMISSIBILITY AND EVIDENTIARY STRENGTH

Organizations typically retain records and other information in order to address business, operational, legal, regulatory, and/or historical needs. Retaining electronic information in a manner that serves these needs can be a complex task, requiring the right combination of technology, controls, and people to ensure that needed information is trustworthy and reliable over the long term. Furthermore, it can be difficult for those charged with retaining electronic information to determine how the requirements of the courts, regulators, and auditors should affect the way that they evaluate, implement, and configure storage and other information systems.

While federal and state laws have generally evolved to allow and even promote the use of electronic documents, email messages, and other forms of electronic information as evidence,<sup>7</sup> by and large it is up to individual organizations to determine how to retain and manage this information in a manner that promotes admissibility, evidentiary strength and satisfies regulatory requirements. Regulators in many major industry sectors have promulgated regulations that address retention and management requirements for electronic records and information.<sup>8</sup> Some of these regulations stipulate specific criteria for the management of electronic information, and they generally make clear that organizations are expected to manage electronic records and data in a manner that ensures its accuracy, completeness, reliability, accessibility, integrity, and trustworthiness.<sup>9</sup> The criteria used to assess Centera in this Evaluation are derived in part from such requirements.

Aside from the more limited set of information typically subject to regulation, nearly any piece of information that is retained by an organization can become evidence in the course of litigation, audits, investigations and other formal proceedings.<sup>10</sup> In determining the admissibility and persuasiveness of electronic evidence, the courts follow procedures and principles that have been established by case law and rules such as the Federal Rules of Civil Procedure, the Federal Rules of Evidence, and the Uniform Rules of Evidence. The criteria used to assess Centera in this Evaluation are derived in part from such rules.

Electronic records and other digital information offered up in court by organizations must generally meet the business records exception to the hearsay rule, which stipulates, among other things, that records must have been “kept in the course of a regularly conducted business,” and “the source of information or the method or circumstances of preparation” must be trustworthy.<sup>11</sup> Evidence then must be authenticated, that is, it must be demonstrated that the evidence is “is what its proponent claims.”<sup>12</sup> Evidence that is not properly authenticated can be excluded, as was the case in Pettiford v. N.C. HHS, where email evidence was excluded, even though it supported the defendant’s case. The authenticity of a record can be challenged in many ways, including: by asserting that that it has been altered;<sup>13</sup> by demonstrating that the system that created or retained the record was not reliable;<sup>14</sup> and by questioning the record’s authorship or provenance.<sup>15</sup> While an untrustworthy electronic record can be excluded outright, in many cases, a successful challenge of an electronic record’s authenticity may not affect its admissibility as evidence, but rather will affect its persuasiveness or weight.<sup>16</sup>

Given the possible challenges to the admissibility and persuasiveness of electronic evidence, organizations need to invest in the technology, people, and controls that will ensure that digital information is retained and managed in a trustworthy fashion.

KAHN CONSULTING, INC.

157 LEONARD WOOD NORTH • HIGHLAND PARK IL • 60035

PHONE: 847.266.0722 • FAX: 847.266.0734 • E-MAIL: INFO@KAHNCONSULTINGINC.COM

---

### III. ABOUT CENTERA

---

EMC's Centera Compliance Edition product ("Centera") is designed to provide a long-term storage solution for "fixed content" – static information that is intended to be retained in a trustworthy manner for a predetermined period of time. Such fixed content includes financial documents, e-mail messages, digital images, and many other types of information that must be kept for business, operational, legal, compliance, and/or historical purposes. EMC designed Centera to provide evidentiary benefits while leveraging the economic and functional benefits of magnetic disk.

Centera is designed to enable the storage of fixed content in a manner that:

- Ensures content integrity, authenticity, security, completeness and accessibility over the long term, in accordance with relevant laws and regulations
- Offers fast, online access to content
- Minimizes the burden of system configuration and management
- Reduces the disruption and expense of media migration
- Supports business continuity and data recovery needs
- Allows storage repositories to grow in a non-disruptive, flexible manner
- Integrates with existing information and records management applications

**Centera Architecture.** Centera is an integrated combination of software and off-the-shelf hardware components sold as a cabinet with an expandable storage capacity. Centera itself is not an information or records management application, but rather an online information repository that works transparently "behind the scenes" to retain, protect, and retrieve the content produced by such applications. Centera can be connected to and integrated with a broad range of software applications ("controlling applications") within multiple markets, including, but not limited to: medical imaging; e-mail archiving; enterprise content management; records management; e-learning; audio and video management; workflow; and so on. To expedite deployment and integration, EMC has an open application programming interface (API) available for partners to integrate Centera with many applications in a variety of industries.

To meet its design goals, Centera incorporates several unique features. This Evaluation focuses on those features designed to meet general criteria for the secure long-term storage of trustworthy records and business information for legal and regulatory purposes.

#### CENTERA FEATURES

##### Content Addressing

EMC's Centera uses a data access paradigm known as "Content Addressing." Content Addressing is a method for storing, accessing, and authenticating a digital file, document, or image (collectively referred to as "digital object" throughout this Evaluation). Content Addressing creates and uses secure alphanumeric object descriptors derived from the content of the object itself. This "content addressing" method differs from traditional "location addressing" methods where digital objects are stored and accessed based on their physical or logical location

KAHN CONSULTING, INC.

157 LEONARD WOOD NORTH • HIGHLAND PARK IL • 60035

PHONE: 847.266.0722 • FAX: 847.266.0734 • E-MAIL: INFO@KAHNCONSULTINGINC.COM

within the storage system. With Content Addressing, the storage and retrieval of content occurs independently of its physical storage location, and no URLs, file structures, or pathnames are used at all. Content Addressing also offers several evidentiary and compliance benefits, as described below.

### **Redundant Array of Independent Nodes (RAIN)**

Each Centera cabinet includes 8 - 32 “nodes” that provide 7.6 - 30.8 terabytes of raw data storage. Multiple Centera racks can be configured as a “cluster” containing up to 180 terabytes of raw storage. Each node is an independent unit comprised of a motherboard, a processor running Centera software, and four magnetic disk drives. A node either functions as a “Storage Node” that serves to store digital objects over the long term, or an “Access Node” that facilitates the interactions between a controlling application and the Storage Nodes. The ability of each interconnected node to process and store data independently reduces dependence on a central system (this providing greater reliability), and supports advanced data protection, replication, and recovery, as described in detail below.

### **Content Protection Mirroring (CPM) and Content Protection Parity (CPP)**

Centera provides two user configurable methods for continuously protecting against the loss of digital objects in the system due to data corruption, device failure and so on. The first method, Content Protection Mirroring (CPM), automatically creates two physical copies of an object on two separate nodes within the system. Using this method, there are always two complete copies of each object within the system; copies that can subsequently be used for data recovery and record regeneration purposes.

The second method, Content Protection Parity (CPP), automatically splits files larger than 20 KB into six data fragments and a seventh “parity” fragment. Each of the seven fragments is stored on separate nodes. If any one node were to fail, the missing fragment stored on the failed node can be automatically recreated using any of the remaining five fragments and the parity fragment.

Data protection schemes such as CPM and CPP support the long-term storage of fixed content, as described below in greater detail.

### **Data Regeneration: “Self Healing”**

Centera continuously monitors each drive and node for faults in either the node, the drive, or in the stored objects themselves. If a fault is detected in a disk or node, that disk or node is isolated and its content is automatically replicated to different nodes using the mirrored copy or parity fragments. Similarly, if corruption is found within an individual object, that object is “regenerated” using the mirrored copy or the parity fragments (depending upon the method originally used to store the file). These operations are completed automatically and without disruption to the overall functioning of the Centera system. However, the system is configured to automatically notify a system administrator when such an operation occurs. The repair and rebuilding of stored data is an important capability of a long-term storage system.

---

## IV. CENTERA CAPABILITIES

---

Electronic business records and other fixed content must be stored and managed in a trustworthy manner. There is little point in expending resources to store content if it cannot be relied upon for business, operational, legal, compliance, and historical purposes. Furthermore, electronic evidence that cannot be properly authenticated may have diminished value as evidence and will likely carry less weight with the courts.

Trustworthiness is most accurately thought of as a quality that results from the sum total of the people, policies, procedures, environments, strategies, and technologies used throughout the lifecycle of a business record. Trustworthiness suggests that a court, regulator, auditor – and the organization itself – can trust and rely upon the content of a record.

The technology used to store and manage digital information plays an integral role in ensuring the trustworthiness of the stored information. As stated in the Federal Rules of Evidence, evidence can be authenticated by “evidence describing a process or system used to produce a result and showing that the process or system produces an accurate result.”<sup>17</sup>

This part of the Evaluation is divided into sections that describe capabilities that are desired in storage systems; explain why each capability is desired; and assess Centera’s compliance with each capability.

### 1. LONG-TERM CONTENT INTEGRITY

**Desired Capability.** Electronic records and business information should be protected from inadvertent or deliberate alteration. A system that protects records from alteration minimizes the likelihood that the authenticity of electronic records will be challenged in court.

**Information Management Principle.** Information has integrity if it can be demonstrated that it has not been altered and remains accurate since it was created or archived. Unlike paper-based information, which has inherent features that deter alteration (such as the physical bond between ink and paper), the alteration of most digital information in its native form is easily accomplished without detection. Business best practices and many laws and regulations require digital information to have integrity.

**Centera Capabilities.** The Centera Content Addressing system works to prevent the inadvertent or deliberate alteration of information, as follows:

- 1) **Hashing.** All information sent to Centera by a controlling application is processed by an MD5 message-digest (or “hashing”) algorithm.<sup>18</sup> This algorithm processes a digital object of any size at the binary level to produce a fixed-length (128 bit) “fingerprint” of the object. This fingerprint is the unique byproduct of that digital object, and that digital object always creates the same fingerprint when it is processed by the algorithm. Consequently, an object’s current fingerprint can be compared to the fingerprint it had when it was first stored in order to determine if the object has changed. A different fingerprint indicates that the document has changed, even if only by a single bit.

KAHN CONSULTING, INC.

157 LEONARD WOOD NORTH • HIGHLAND PARK IL • 60035

PHONE: 847.266.0722 • FAX: 847.266.0734 • E-MAIL: INFO@KAHNCONSULTINGINC.COM

- 2) **Content Addressing.** The object's fingerprint is used by Centera as that object's Content Address (CA). The CA is unique to that object and is stored and used by Centera to access and authenticate the object throughout its entire lifecycle.
- 3) **Validation.** Each digital object's CA is recalculated during all significant interactions between Centera and the controlling application, and is also continuously calculated by built-in Centera data validation utilities that run perpetually within each node, comparing each object's current CA to its original CA.
- 4) **Deliberate Alteration.** There is no direct access to the files in Centera, as described in detail below. Files can only be altered within a controlling application after retrieving the object from Centera. When a user saves the altered object back into Centera, Centera will recognize that the object has been altered, and will treat it as an entirely new object with a new Content Address. In this manner, Centera works to prevent the deliberate alteration of objects stored within it.<sup>19</sup>
- 5) **Inadvertent Alteration.** If an object is corrupted or otherwise altered inadvertently, this change will be revealed by the automatic, ongoing comparison of CAs, and the altered file will be automatically replaced with the mirrored file, or rebuilt from the parity fragments. EMC calls this process "organic regeneration." In this way, Centera protects against inadvertent alteration of stored data.

## 2. CONTENT COMPLETENESS AND AUTHENTICITY

**Desired Capability.** Information storage systems should retain electronic records in manner that preserves their complete content, physical form, layout, and metadata – especially that metadata indicating origin and provenance.

**Information Management Principle.** An electronic record is said to be authentic if it is in fact "what it purports to be." That is, the source or origin of the e-record can be reliably demonstrated. This often requires proof of who generated an e-record, and who controlled it at a certain time. In addition, an electronic record should remain in a complete and accurate form throughout its lifecycle in order to be considered trustworthy, and to satisfy a variety of business and legal requirements.

**Centera Capabilities.** Information sent to Centera by a controlling application is retained by Centera in a manner that preserves the object's original qualities, such as the file format, physical appearance, content, and metadata. Changes to an object as small as a single bit can be detected and corrected by Centera through the use of the CA. The ability to detect such changes is maintained throughout the entire time that a file is stored within the system. In this manner, Centera provides features that enable electronic records to be stored in a complete and authentic manner.

KAHN CONSULTING, INC.

157 LEONARD WOOD NORTH • HIGHLAND PARK IL • 60035

PHONE: 847.266.0722 • FAX: 847.266.0734 • E-MAIL: INFO@KAHNCONSULTINGINC.COM

### 3. CONTENT ACCESSIBILITY

**Desired Capability.** Organizations should be able to access information in a timely, trustworthy, and cost-effective fashion at any time during the information lifecycle. As stated by the courts, “[u]tilizing a system of record keeping which conceals rather than discloses or makes it unduly difficult to locate” may be considered the equivalent of destroying records.<sup>20</sup>

**Information Management Principle.** Information that cannot be readily found and accessed is of marginal utility. In the short term, responding to a regulator or a court request for records must often be completed in a short timeframe. In the long term, numerous factors such as the limited lifespan of every digital storage medium; data corruption; heterogeneous storage formats; technological obsolescence; and other factors threaten the long-term accessibility of electronic information and records.

**Centera Capabilities.** Centera addresses content accessibility issues as follows:

- 1) **Short-term accessibility.** Unlike solutions based on optical disk, tape, and other storage formats typically used for the archiving of electronic records and other fixed content, Centera uses magnetic disk. Magnetic disk can provide faster access times than certain other media. In addition, the Centera architecture enables an entire archive of information to remain “online” and accessible without significant performance degradation. Conversely, systems that rely on removable media, such as optical disk, typically employ a staged system where only a certain number of disks remain in the storage device for immediate access. In this regard, Centera may provide faster and more cost effective short-term access to information than other kinds of storage systems designed for fixed content.
- 2) **Metadata and Indexing.** When the controlling application stores an object in Centera, it also creates and stores an Extensible Markup Language (XML) file containing the CA and metadata about the object. This XML file is known as a Content Descriptor File (CDF), and can contain both “standard” information such as filename and a time-date stamp, as well as “custom” metadata stipulated by the controlling application, such as a project name or office number, for example. The data in the CDF can subsequently be used for querying purposes, controlling retention periods, and for other purposes.
- 3) **Media Aging and Migration.** The health of each disk drive within Centera is continuously monitored. Disk and node failures that can be “self-healed” are repaired automatically by isolating the failed node and recovering its data onto other nodes in the Centera cluster, at which stage the failed node can be replaced by an administrator.<sup>21</sup> This capability effectively results in the ongoing migration of data from aged to fresh media. In addition, Centera’s design is hardware-independent, which allows it to adapt to the latest disk drive technology. This important capability may help to ease future upgrades or migrations to different drives as new technology is developed.

KAHN CONSULTING, INC.

157 LEONARD WOOD NORTH • HIGHLAND PARK IL • 60035

PHONE: 847.266.0722 • FAX: 847.266.0734 • E-MAIL: INFO@KAHNCONSULTINGINC.COM

#### 4. INFORMATION SECURITY

**Desired Capability.** Storage systems should provide information security controls and capabilities that protect the system and its contents from alteration, corruption, inaccessibility, loss, compromise of confidentiality and privacy, and other events.

**Information Management Principle.** Organizations manage and store valuable information that must be protected. In some cases, confidentiality must be maintained, and in other cases privacy protection is a legal requirement. Security is a complex process that involves many different procedures and technologies, but it is fundamental to an organization meeting its information management goals and obligations.

**Centera Capabilities.** Centera offers a variety of controls and techniques designed to secure the system and its contents, as follows:

- 1) **Architecture.** Centera is not a “browseable” or directly-accessible system. The only access to Centera is through the controlling application or an administrative console. This architecture makes it difficult for an attacker without access to either of these entry points to find, view or access content within a Centera cluster.<sup>22</sup>
- 2) **Access controls.** The controlling application’s access to Centera can be policed in a variety of ways at the Centera System Administration level, including password protection and through configurable file operation (i.e., query, delete, retrieve, and store) protection.
- 3) **Administration.** Centera can be configured to disallow remote administrative access. This would limit administrative access to individuals who have a direct physical connection to a Centera cabinet, which typically would be located in a physically secure location.
- 4) **Application Access.** A controlling application’s access to Centera is strictly limited to designated Internet Protocol (IP) addresses and port numbers.

#### 5. RESISTANCE TO DELETION AND OVERWRITING

**Capability.** Storage systems designed to store highly sensitive and regulated information should offer the capability to protect information from being inadvertently or deliberately deleted or overwritten. Improper deletion or “spoliation” of evidence can lead to serious consequences inside and outside the courtroom. In litigation, the penalties can include severe fines and penalties, and even the overall dismissal of a claim.<sup>23</sup>

**Information Management Principle.** In order to satisfy certain business requirements, laws, regulations and other criteria, electronic records may need to be stored in a fashion that ensures that they cannot be deleted or overwritten. From an evidentiary perspective, such a capability helps to demonstrate record integrity and preempt attacks on record trustworthiness.

KAHN CONSULTING, INC.

157 LEONARD WOOD NORTH • HIGHLAND PARK IL • 60035

PHONE: 847.266.0722 • FAX: 847.266.0734 • E-MAIL: INFO@KAHNCONSULTINGINC.COM

**Centera Capabilities.** A controlling application can stipulate the period of time that an object sent to Centera must be retained. This information is contained within the CDF metadata, and is associated with the object throughout its lifecycle. Once an object has been designated in this manner, the object and its associated CDF cannot be deleted or overwritten before the expiration of the retention period. Furthermore, once the retention period has expired, Centera does not proactively delete expired content. Rather, deletion must be initiated by the controlling application.

Controlling applications communicate with Centera through a programmatic interface known as an Application Programmer's Interface (API). The API allows only a predetermined set of Centera functions to be executed by the controlling application. Specifically, the API provides access to five basic functions, one of which is the "delete" command. However, the Centera software is written so that a "delete" command cannot be executed on an object that has an unexpired retention period stipulated in its CDF.

In this manner, Centera preempts the unauthorized deletion or overwriting of stored information.

## 6. RECORDS RETENTION

**Capability.** A storage system designed for the long-term storage of electronic records should offer records retention functionality.

**Information Management Principle.** Laws, regulations, standards, and practices require organizations to retain specific types of information for specified periods of time. Organizations retaining records in electronic form require storage systems that enable them to designate retention periods for electronic records and dispose of records at the end of their lifecycle.

**Centera Capabilities.** Centera enables controlling applications to designate record retention periods. This information is stored in the CDF file that is associated with object throughout its lifecycle and protects the object from being deleted or overwritten before the end of its retention period. After an object is stored in Centera, its retention period cannot be shortened. It can however, be lengthened if changing retention criteria, legal hold requirements, or other factors necessitate an extension of the original retention period.

Default retention periods can also be applied to objects. In the case where a controlling application does not specify a retention period, Centera can be configured to automatically assign a retention period of 0, in which case the file could be disposed of at any time; or to automatically assign an "infinite" retention period, which would prevent the deletion of that object from Centera at any time in the future. These default retention periods can also be set by the controlling application.

## 7. BUSINESS CONTINUITY AND DISASTER RECOVERY

**Desired Capability.** Standard disaster recovery techniques require that data is stored in at least two physically separate locations. This is also a requirement of some regulations, such as SEC

Rule 17a-4(f)(3)(iii), which requires that securities firms “[s]tore separately from the original, a duplicate copy of the record . . .for the time required.”

**Information Management Principle.** Data that does not exist in two or more places can be permanently lost if the device or facility housing the data is damaged or destroyed. Thus, there is a need for organizations to copy important data to different physical locations for backup, disaster recovery, and business continuity purposes.

**Centera Capabilities.** Centera software can be configured to continuously and asynchronously replicate the contents of one Centera installation or “cluster” to a physically separate Centera cluster. This capability will aid an organization in meeting its business continuity and disaster recovery needs as they relate to information stored within Centera. In addition, each Centera cabinet can be powered by two independent sources of AC power, another capability that supports business continuity requirements.

## 8. RECORDING AND STORAGE PROCESS INTEGRITY AND AUTHENTICITY

**Desired Capability.** When organizations archive electronic records and business information for future use, the reliability and integrity of the initial recording and storage process should be validated.

**Information Management Principle.** Information cannot be relied upon unless there is assurance that the information was recorded in a manner that reflects the form and content of the information as it was originally created.

**Centera Capabilities.** The CA of a digital object is calculated *before* it is written to disk within Centera, and also immediately thereafter. Next, the two CAs are automatically compared to detect any changes that may have occurred during the recording process. This operation ensures that the object that is stored within Centera is the same object that was sent to Centera by the controlling application.

## 9. RECORDS DISPOSITION

**Desired Capability.** Records Management solutions should provide the capability to properly dispose of information once it is no longer needed.

**Information Management Principle.** Disposition is the final lifecycle stage of most information. Although there are relatively rare cases where information must be retained in perpetuity for historical and other purposes, over time the vast majority of information ceases to be of value to an organization. In the digital world, it can be difficult and expensive to ensure that electronic information is properly disposed of. This can lead to situations where files are not properly disposed of and unwanted files are recovered or recreated in the course of litigation.

**Centera Capabilities.** When Centera receives a request from a controlling application to delete an object, Centera first determines if the object has a retention period designated within its CDF. If the retention period is still valid, then the object cannot be deleted. If the retention period has expired, the object is now eligible to be deleted. Consequently, upon receiving a delete command from the controlling application, Centera will delete the object and its CDF, and automatically recover the disk space for further use.

KAHN CONSULTING, INC.

157 LEONARD WOOD NORTH • HIGHLAND PARK IL • 60035

PHONE: 847.266.0722 • FAX: 847.266.0734 • E-MAIL: INFO@KAHNCONSULTINGINC.COM

In addition, for added certainty in the disposition of objects, Centera can be configured to automatically use digital “shredding” techniques that conform to the US Department of Defense 5220.22-M standard for permanently deleting digital information.

It should be noted that a feature of Content Addressing is that identical objects are never stored more than once within Centera. For example, if the controlling application submits an e-mail message for archiving to Centera, and a calculation of that message’s CA reveals that it is identical to a message already stored within Centera, the message will not be archived again (which, among other things, promotes system storage efficiencies). Rather, a new CDF will be created that “points to” the original stored e-mail message, rather than storing an additional, identical copy. Consequently, multiple CDFs can point to the same object, and the object would continue to exist until the CDF with the longest retention has expired and has been deleted.<sup>24</sup> Among other things, this architecture supports situations where a single object is subject to multiple, different retention periods due to regulation or policy, or when the retention period for an object needs to be extended due to investigations, audits, or litigation.

KAHN CONSULTING, INC.

157 LEONARD WOOD NORTH • HIGHLAND PARK IL • 60035

PHONE: 847.266.0722 • FAX: 847.266.0734 • E-MAIL: INFO@KAHNCONSULTINGINC.COM

---

## VI. ABOUT KAHN CONSULTING

---

Kahn Consulting, Inc. (“KCI”), headed by founder and principal Randolph A. Kahn, Esq., is a consulting firm specializing in the legal and policy issues of information technology. Through a range of services including information management program development, risk management audits, policy development and evaluation, product assessments, legal and compliance research, and education and training, KCI helps its clients address today’s critical issues in an ever-changing regulatory and technological environment. Based in Chicago, KCI provides its services to Fortune 500 companies and state and federal governmental agencies in North America and around the world. Kahn has advised clients such as McDonalds Corp., Hewlett-Packard, United Health Group, the Federal Reserve Banks, Ameritech/SBC Communications, Motorola, Mutual of Omaha, and the Environmental Protection Agency. More information about KCI, its services and its clients can be found online at: [www.KahnConsultingInc.com](http://www.KahnConsultingInc.com).

---

<sup>1</sup> In undertaking this engagement, KCI exclusively relied upon information supplied by EMC through internal and external documentation, and interviews with EMC representatives, including senior system designers. KCI does not conduct independent laboratory testing of information technology products, and as such, did not evaluate Centera in a laboratory setting or otherwise field-test any EMC products.

<sup>2</sup> Lyman, Peter and Hal R. Varian, "How Much Information," 2000. Online: <http://www.sims.berkeley.edu/how-much-info> (Hereafter referenced as “Lyman”).

<sup>3</sup> Throughout this document, the term “organization” is used to refer to companies, businesses, government agencies, institutions, associations, and other entities conducting profit and not-for-profit activities.

<sup>4</sup> Lyman.

<sup>5</sup> IDC Report, “Worldwide E-mail Usage Forecast, 2002-2006: Know What's Coming Your Way,” as reported by Gretel Johnston, IDG News Service, October 2, 2002.

<sup>6</sup> *Bills v. Kennecott Corp.*, 108 F.R.D. 459, 462 (D. Utah 1985).

<sup>7</sup> See, for example, electronic record-enabling laws such as the federal Electronic Signatures in Global and National Commerce Act (E-SIGN), the state Uniform Electronic Transactions Act (UETA), and the Government Paperwork Elimination Act (GPEA), which applies to federal government agencies.

<sup>8</sup> Examples of such regulators, and their corresponding regulations include: the Securities & Exchange Commission (17 CFR § 240.17a-4), electronic records; the Environmental Protection Agency (55 Fed. Reg. 31,030 (1990)), policy on electronic reporting; the Food and Drug Administration (21 C.F.R. Part 11), electronic signatures and records; the Internal Revenue Service (Treasury Reg. 301.6061-1), signature alternatives for tax filings; and the Commodity Futures Trading Commission (17 C.F.R. Part 1.4 and Part 1.3(tt)), electronic signatures for filings.

<sup>9</sup> See, for example the requirement for “non-rewriteable, non-erasable” electronic storage media in SEC Rule 17a-4(f)(2)(ii)(A).

---

<sup>10</sup> See, for example, FED. RUL. CIV. PROC. 34(a), “Any party may serve on any other party a request . . . to produce . . . to inspect and copy, any designated documents (including writings, drawings, graphs, charts, photographs, phonorecords, and other data compilations from which information can be obtained, translated, if necessary, by the respondent through detection devices into reasonably usable form).” Also, FED. RUL. CIV. PROC. 26(b)(1): “Parties may obtain discovery regarding any matter, not privileged, that is relevant to the claim or defense of any party, including the existence, description, nature, custody, condition, and location of any books, documents, or other tangible things and the identity and location of persons having knowledge of any discoverable matter. . . . Relevant information need not be admissible at the trial if the discovery appears reasonably calculated to lead to the discovery of admissible evidence.”

<sup>11</sup> Federal Rules of Evidence (hereafter “FRE”) 803(6).

<sup>12</sup> FRE 901(a).

<sup>13</sup> See, for example, *United States v. Whitaker*, 127 F.3d 595, 602 (7th Cir. 1997).

<sup>14</sup> See, for example, *United States v. Dioguardi*, 428 F.2d 1033, 1038 (2d Cir. 1970).

<sup>15</sup> See, for example, *United States v. Simpson*, 152 F.3d 1241 (10th Cir. 1998).

<sup>16</sup> See, for example, *United States v. Catabran*, 836 F.2d 453, 458 (9th Cir. 1988).

<sup>17</sup> FRE 901(b)(9).

<sup>18</sup> The MD5 algorithm is widely considered to be an industry standard.

<sup>19</sup> While a skilled attacker can circumvent even the strictest security controls within any information system, given enough knowledge, resources, and time, EMC had the Centera product reviewed by *Internet Security Systems* and *@Stake*, both of which concluded that the product was well protected.

<sup>20</sup> See, for example, *Kozlowski v. Sears Roebuck & Co.*, 73 F.R.D. 73 (D.Mass.1976).

<sup>21</sup> No single storage subsystem is immune from data loss if several hardware components fail simultaneously. EMC recommends that organizations employ the disaster recovery/replication features of Centera, which help to minimize the likelihood that data will be lost due to node failures or catastrophic events.

<sup>22</sup> While a skilled attacker can circumvent even the strictest security controls and mechanisms within any information system, given enough knowledge, resources, and time, Centera has built in substantial features to prohibit such an attack and minimize any resulting harm to stored content. See also footnote 19.

<sup>23</sup> See, for example, *Mathias v. Jacobs* 197 F.R.D. 29, S.D.N.Y., 2000.

<sup>24</sup> Because Centera allows multiple CDFs with different retention periods to point to the same object, there will be cases where one individual no longer has access to a certain object because they have deleted “their” CDF, while another individual still has access to the object because “their” CDF still exists. When making representations to courts about the availability of information in the context of discovery proceedings, organizations should be aware of this feature of Centera and address its proper use through policy, procedure, and controlling application configuration.