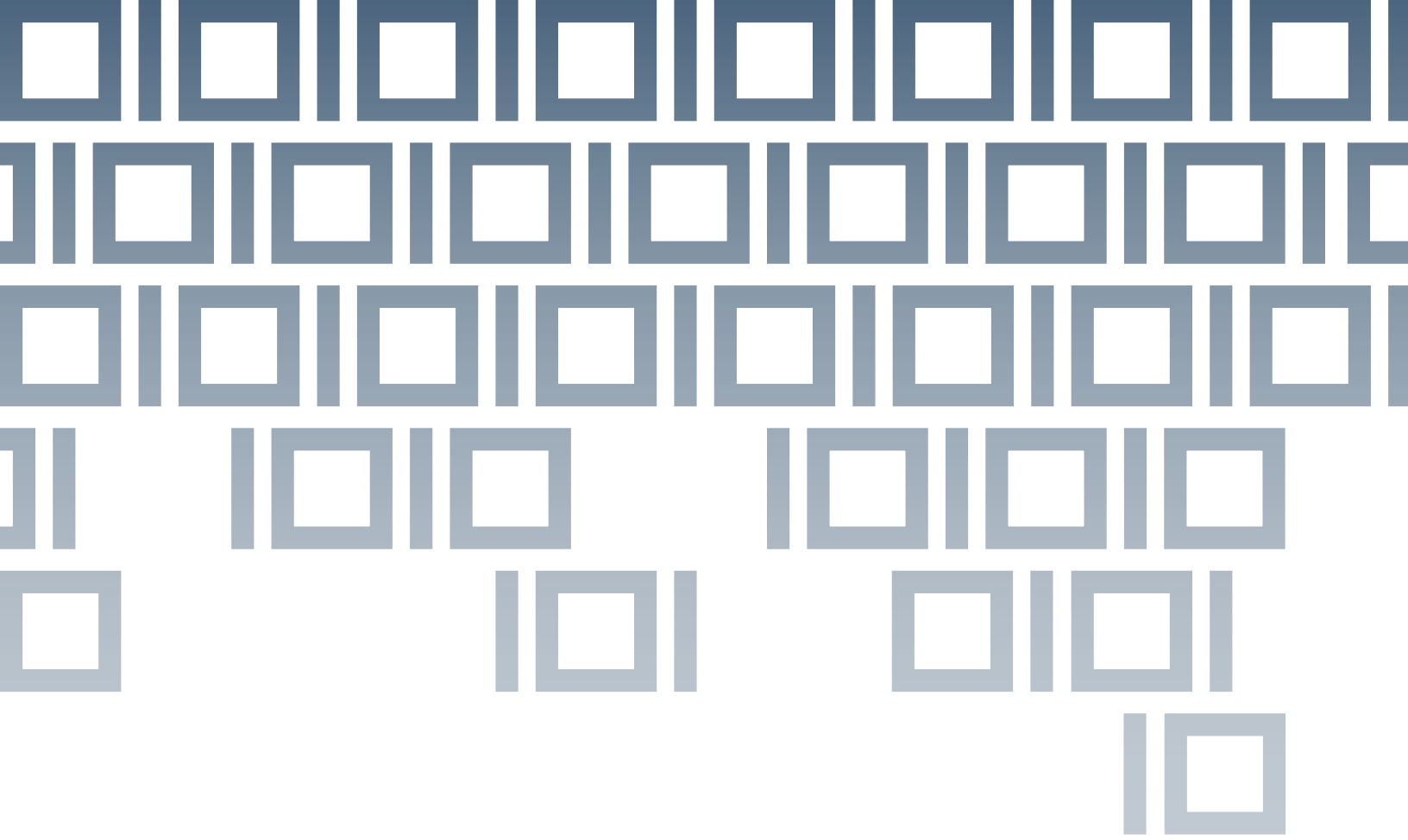

# 7 Essential Steps for Taking Control of Digital Data Debris

KAHN CONSULTING INC.



Sponsored By OPENTEXT

# Introduction

## Organizations across all industry sectors are attempting to control the mounting flood of digital information being generated daily—90% of it unstructured.[1]

"Google processes more than 24 petabytes of data per day, a volume that is thousands of times the quantity of all printed material in the U.S. Library of Congress."

—*Big Data,* by Viktor Mayer-Schonberger and Kenneth Cukier

But Information Governance in 2014 and beyond is not just about dealing with information volumes. It's about understanding what information exists, where it exists, making it accessible, and managing it. What information needs to be retained to satisfy laws? What has business value? And can the final contract or client file be found in time? How can the organization operate "faster, better, cheaper," and be legally compliant when so much digital data debris exists?

Though we are well into a still-maturing information-intensive economy, what is less evolved is the way most businesses think about information. Simply stated, most believe the more the merrier. That is, the more information they have, the better off they are. Yet we are clearly in a period of diminishing returns for information accumulation. **Today, finding the right field, to find the right haystack, in a miraculous attempt to find that one required needle is nothing short of monumental.**

Further, many business professionals believe that storage is cheap. This is an all-too-often cited mantra that is fallacious at best. While storage cost per terabyte may be going down by a few percent annually, that decline is dwarfed by a 20%-50% information footprint growth rate.[2] There are companies that spend tens, even hundreds of millions of dollars just to store information, which is a recurring cost. The smart decision is to clean the corporate information dumping grounds of their dead data. It just makes business sense.

[1]IDC White Paper: The Knowledge Quotient: Unlocking the Hidden Value of Information Using Search and Content Analytics; June 2014
[2]InformationWeek 2011 State Of Storage Survey

# Welcome to the Era of the Law of Diminishing Returns

A widely accepted business principal states that information is an asset, but at some point, too much information becomes a liability and quickly loses its value. Part of the decline is due to an expanding universe of communication and collaboration technologies that create a lot of short-term (transitory) information with limited longer-term business or legal value. The decline in value is also driven by the challenge businesses have finding data they need quickly in the mountain of information on which they are sitting.

According to the Council for Information Auto-Classification's survey, *The Information Explosion* (InfoAutoClassification.org), enterprises now have so much information that **nearly 50% of companies need to recreate business information because they can't find the original.**

Organizations are also facing increasing regulatory pressure, enforcement, and public scrutiny on all of their data storage activities. When combined with growing data volumes, the issues of information privacy, security, protection of trade secrets, and records compliance become more complex and high risk. Those issues include:

- Storage costs have sharply increased, with some companies refusing to allocate any more storage to the business. The user reaction, out of necessity, is to store their data wherever they can find a place for it. (Don't buy the argument that storage is cheap—everyone is spending more on storing unnecessary data, even if the per-gigabyte media cost has gone down).

- Litigation and discovery costs are soaring as organizations have lost track of what is where, who owns it, and how to collect, sort, and process it.

- Buried intellectual property, trade secrets, personally identifiable information (PII), and regulated content are subject to leakage and unauthorized deletion, and are clear targets for opposing counsel—or anyone who can access them.

- There is a lack of centralized policies and systems for the storage of records results in hard-to-manage "record sites" spread throughout the organization. (Many companies have hundreds or even thousands of rogue SharePoint repositories).

- There is a lack of a clear strategy for managing records that have long-term, rather than short-term, business, legal, and research value, as well as a lack of a policy for managing content that has only short-term value (i.e. non-records).

**88%** of organizations surveyed have no idea of the content in their stored data.

**58%** of organizations are keeping information indefinitely.

**79%** of organizations say too much time and effort is spent manually searching and disposing information that has met its retention requirements.

**58%** of organizations still rely on employees to decide how to apply corporate policies.

Source: The Information Explosion survey from the Council for Information Auto-Classification http:/infoautoclassification.org/

# What Is Defensible Disposition and How Can It Help?

One solution to the unmitigated data sprawl is to defensibly dispose of business content that no longer has business or legal value to the organization. In the old days of Records Management, it was clear that courts and regulators alike understood that records came into being and eventually were destroyed in the ordinary course of business. It is considered good business practice to destroy unneeded content, provided that the rules upon which those decisions are made consider legal requirements and business needs. The good business practice of "cleaning house" of old records has become taboo for some businesses. Now it needs to start again.

Defensibly disposing of unneeded information helps an organization achieve a thinner information footprint by doing what Records Management does—cleaning house of information that is no longer needed. The difference now is that the amount of information and number of places it is located make it more challenging (and costly) to undertake. The new business information environment with terabytes and petabytes of unmanaged content requires Defensible Disposition processes, which utilize automation instead of a manual document-by-document method, that relies on people to do the heavy lifting. Very often, information analysis is done by analytics and/or classification technologies because employees can't reliably review hundreds of millions of files.

While there is no single approach that is right for every organization, records and legal teams need to take an informed approach, looking at corporate culture, risk tolerance, and litigation profiles. Building a Defensible Disposition process with technology, policies, procedures, and management controls designed to ensure that records are managed over time and properly disposed at the end of their lifecycle is essential for organizations with lots of outdated information.

Here are the Seven Essential Steps for your organization to take control of your digital data debris:

## STEP 1: ASK FOR EXECUTIVE SUPPORT

To get information management right takes three things. First, you need the right technologies including data capture, data classification, data management, data security and sharing, storage repositories, and disposition capabilities—all of which can be found in an enterprise content management (ECM) system.

Second, it takes experienced personnel to ensure information management practices are set up, configured (including policies that are legally defensible), and maintained properly.

Third, it takes a management team with varying backgrounds and areas of responsibilities to create an organization-wide IM strategy and ensure it's audited and followed.

Rallying the troops to make Defensible Disposition happen will also require executive support and leadership from at least two corner offices: the CIO and General Counsel. This is because the complexity, cost, and enterprise transformation will need to be jointly supported, messaged, and advanced by senior IT and legal leadership. Large quantities of information will be disposed of, which will end up producing a (usually) large net cost savings as well as positively impact legal processes and responsibilities. All of these factors need to be addressed up front. If you can't get IT and legal on board, Defensible Disposition won't happen—so start at the top.

Get executive support from both the CIO and GC, then build your multidisciplinary information management program team to make Defensible Disposition come to life. This multidisciplinary team should consist of high-level representatives from groups such as:

- IT
- Legal
- Records Management
- Information Security
- Marketing
- Compliance
- Risk Management
- HR
- Tax and Audit
- Finance

## STEP 2: BUILD A CONSERVATIVE BUSINESS CASE

Cleaning out your digital data debris will take time and money. But the advantages of running a leaner information footprint include a myriad of business and risk-reduction benefits.

To get executives interested and on board with cleaning up old information, you will need to advance a compelling business case. While there are many soft cost savings, business benefits, and risk reductions that make a great story for taking on the task, the most compelling business case is a conservative one—based on hard costs that are easily quantifiable. In that regard, unearthing storage costs and potential savings for the various target data repositories is a great place to start. If you have a large footprint that costs five million to ten million dollars per petabyte per year to store information, you can quickly see that reducing the footprint by 10%, 20%, or 30% can have big economic savings. Rather than float questionable soft numbers, build a business case that demonstrates real numbers of both savings as well as costs to convince the higher-ups to undertake the initiative.

## STEP 3: DEVELOP A DILIGENCE PROCESS BASED ON DATA CONTENT

To defensibly rid the organization of unneeded information, you need to address two key issues before any disposition can commence. If the information has any ongoing business value, then it shouldn't be destroyed. And even if information is no longer needed for business purposes, if the information needs to be preserved for an audit, investigation, or lawsuit, then it can't be destroyed without legal risk. Addressing those two issues when dealing with large volumes of information, each of which is not manually reviewed, is integral to Defensible Disposition.

This is why Defensible Disposition is considered different than Records Management. Though Records Management includes retention rules which allow the disposal of expired records to occur, Defensible Disposition goes beyond records to address all unneeded or valueless content within a repository or enterprise, not just content considered a record. The only way to know whether you can dispose of large volumes of data is to apply a reasonable diligence process. In the old days, it was considered good

business practice to destroy unneeded content, provided that the rules upon which those decisions were made considered legal requirements and business needs, and were applied in a consistent manner. So, what is a reasonable and good business practice when faced with large quantities of aging electronic data?

The short "lawyerly" answer is, "it depends." Not all target data repositories are the same and therefore the diligence required will vary. For example, if there are thousands of old backup tapes from 30 years ago that haven't been used for years and that no investigation or lawsuits impact, then opening up the tapes just to see what's inside may not be needed before destruction can happen. On the other hand, if you want to clean up the contents of the shared drives, then it's ill-advised to take wholesale action to purge. Instead, establish a process to dig deeper into the information to determine what is needed for retention as well as preservation.

In such situations, the diligence process may rely heavily on technology to evaluate content. For every information repository, the diligence process will be different. But in the end no purging should happen until your organization's lawyers and business stakeholders are comfortable with the analysis performed and legal defensibility of the diligence process.

## STEP 4: EVALUATE INFORMATION REPOSITORIES THAT ARE RIPE FOR CLEANUP

When considering where to start your defensible disposition efforts, try to understand where information hoarding takes place. Evaluate both structured (database) systems as well as unstructured (office automation and text content) environments to see where your company can get the biggest benefit from cleaning house.

Which environments are overrun with ill-managed information (duplicates, useless revisions, expired content, non-business-related files, such as, personal files, etc.)? What dumping ground is costly to the enterprise, given storage volumes and tiers? Which environments create risk, liability, and/or exposure? Can useless information or duplicates be addressed easily? Evaluate the environments for access and use, always considering whether the information may be needed for litigation response.

While unstructured data growth and more specifically, email risk may compel Defensible Disposition, structured environments may be ripe for cleanup as well. Being strategic about where to start an early victory will help you get additional budget and approval to carry on with the project.

## STEP 5: ATTACK THE LOW-HANGING FRUIT

After evaluating the information repositories and environments, the best place to start your attack is the place that allows the easiest cleanup with the least effort, requires the smallest investment, and produces the most value. Don't go after too many environments contemporaneously as it will be overwhelming and raise the risk of failure.

The best place to start could very well be a business unit that is experiencing above average information buildup or data chaos and therefore, would likely benefit the most. It might be that cleaning house in that business unit will reduce information security or privacy risk. Or maybe it's a specific company system that poses the biggest information mismanagement or litigation cost issues for the organization, i.e. the email system or SharePoint infrastructure. In any event, choose a target that will give you a quick, obvious win.

As a general rule, file shares, legacy email systems, and other large repositories will prove to be valuable environments to target, while better-managed document management, records management, and

archival systems may be in less need of remediation. A good time to undertake a cleanup exercise is when your organization doesn't anticipate or isn't in the midst of litigation.

When evaluating potential targets, keep in mind the estimated value received may be reduced storage costs, reduced risks in eDiscovery and regulatory response, increased information security, and/or rising employee productivity, to name just a few.

There is no one right place to start, but understanding these issues will help determine the "low-hanging fruit" in your organization.

## STEP 6: DEVELOP A PROCESS TO ADDRESS RETENTION AND LITIGATION PRESERVATION RESPONSIBILITY

The main reason most organizations are over-retaining and over-preserving is fear. Lawyers have become afraid of doing the wrong thing by getting rid of information that shouldn't have been destroyed. This understandable but unnecessary trepidation stems from the complexity, costs, and risks of eDiscovery. Instead of fixing the problem, they compound it by directing that too much information be preserved or that retention rules be ignored, thereby preempting the routine operation of the records program.

In essence, data repository size and system complexity have driven companies into litigation hell, and now lawyers are making it worse by doing more of the same. Fear leads to bad decisions that don't meet the needs of your business or ironically, keep you out of legal trouble.

To get rid of information, you need to satisfy two legal requirements: 1) If the information is a record, has it met its retention period; and 2) Is the information needed as potential evidence in any anticipated or current legal matter? If the process you have built addresses both questions satisfactorily, you are free to dispose of the information. If you have a records program, failing to dispose of information when the rules dictate undermines your entire program.

Your Defensible Disposition methodology and process should address retention and preservation before you take action to destroy information. What's needed to address these issues will vary, but again, get your lawyers involved early and often. They are your personal insurance policy, as you can tell any judge that your corporate attorney OK'd the disposal.

## STEP 7: USE TECHNOLOGY TO DO THE HEAVY LIFTING

Due to the sheer volumes of data (and its truly exponential growth rate), relying on employees to manage millions of files manually is untenable and wasteful. No executives will ask that employees eschew their regular jobs in favor of reviewing, classifying, and disposing old information.

Trying to conduct a manual, comprehensive, document-level inventory and disposition is neither reasonable nor practical, and will in most cases create limited and poor results. In most instances, Defensible Disposition will necessitate that analytics and/or classification technologies be used to make disposition decisions.

Technology can help discern the meaning of retained unstructured content, in a variety of formats, regardless of where it is stored—and then automate the processes of analysis, classification, retention, and immediate disposal, if appropriate. It is clear that technology solutions, if used correctly, are faster and have higher accuracy levels than people when it comes to classifying information. Increasingly, various studies and the courts make clear that, when appropriate, companies should not fear using technology to help manage information. Relying on technology to do the heavy lifting is not only allowed, but is really the only reasonable way.

# Conclusion

Having information, but not being able to find it, is equivalent to not having it at all. Most businesses with big information piles are at an inflection point: Continue to double the amount of information every two to three years, and within the decade you will cease to be an efficient business. More is not only *not* merrier, it is a drag on costs and efficiency.

A thinner, more streamlined information footprint is essential to continued business success—and the only way to get there is with a comprehensive Defensible Disposition strategy. Having employees destroy a few emails and Word documents every day to begin cleaning up the company infrastructure won't make a dent when you are dealing with terabytes and petabytes of data.

Organizations are not under any obligation to indefinitely retain every piece of information they generate in the course of business. Indeed, in the past, Records Management was your license to clean house of expired records in a legally defensible way. Defensible Disposition simply picks up where Records Management leaves off to help clean up the bigger non-record data collections in a more efficient, legally defensible way. Now get busy.

# About Kahn Consulting, Inc.

Kahn Consulting, Inc. (KCI) is a consultancy specializing in the governance, compliance, and technology issues of information. Through a range of services including information governance strategy and framework development, defensible disposition and information migration strategies, records management program development; electronic records, e-communication and social network policy development; information management compliance audits; technology product assessments; training program development, KCI helps its global clients address today's critical issues in an ever-changing regulatory and technological environment. More information about KCI, its services and its clients can be found online at: www.KahnConsultingInc.com.

KAHN CONSULTING INC.
847.266.0722 • info@KahnConsultingInc.com

Sponsored By
OPENTEXT