



How to Win the Compliance Battle Using

# 'Big Buckets'

Despite verbal skirmishing about how big retention schedule "buckets" should be, RIM professionals agree that reducing the number of records series buckets can help organizations win the battle to increase user accuracy and compliance and reduce costs and risks

**Susan Cisco, Ph.D., CRM**

**A** retention schedule is a formal policy defining an organization's legal, regulatory, and business requirements for records retention and destruction. Applying the big bucket approach when developing or updating retention schedules results in significantly fewer record series or "buckets" and improves the ability of a user or an automatic classification tool to accurately and consistently classify recorded information for retention purposes.

Not everyone agrees that reducing the number of record series or buckets is a prudent approach to managing retention schedules, and lively discussions have ensued at conferences and on the records management listserv. For example, Chris Flynn's big bucket theory blog contains this passage:

*Today we are looking at the 'bucket' and trying to reconcile our traditional methods with a theory that has reached fruition in an alternative discipline ... a schism that is reaching a level of maturation in our profession. In one camp [are] the records managers that hold to traditional time tested and well established practices. In the other camp there are records managers that have leapt on the IT train. They are the prophets of radical change and the canonization of their practices.*

This article explores the controversy around the topic of fewer, bigger retention buckets by clarifying the terminology used to discuss the topic, suggesting the optimum number of buckets in a retention schedule, and explaining how to apply the big bucket approach when developing a new schedule or revising an existing schedule. (Steps for applying the big bucket approach to new or existing retention schedules are in the sidebar article that begins on p. 32.)

The terms "information," "records," "content," and "documents" are used

interchangeably throughout the article, and no attempt has been made to distinguish one from another because the definitions vary so widely by discipline, country, and culture. In the end, all of an organization's recorded information must be managed efficiently and effectively to leverage its value and minimize the risk of retaining it too long or not long enough.

Many organizations have applied controls to the management of their "official records" (records that are legally recognized as establishing some fact and retained as evidence of business transactions or decisions). However, according to Elliot Gerard's presentation at a Microsoft Office Sharepoint Conference in March, the official records of an organization may constitute as little as 5 percent of its total volume of recorded information, leaving 95 percent to be managed using traditional rules and tools.

### **The Big Bucket Approach**

The U.S. National Archives and Records Administration (NARA) is credited with being an early promoter of the "bigger bucket" approach for streamlining granular (usually departmental) retention schedules because the existing processes required federal agencies to commit too much effort to inventorying and scheduling routine records. Initially called "flexible scheduling," the approach provided flexibility in defining record groupings composed of multiple records series in large aggregation retention schedules. Record series that were related to the same or similar business process and were supported by the same or similar legal and business requirements were aggregated into bigger buckets.

Private sector organizations also have found it necessary to streamline their retention schedules for effective implementation of enterprise content management, e-mail archiving, and other recordkeeping systems to ease the classification burden on users, who were having difficulty selecting correctly from

## **At the Core**

### **This article**

- ▶ Describes the controversy around applying the big bucket approach
- ▶ Contrasts the benefits of using big buckets with its potential drawbacks
- ▶ Tells how to apply the big bucket approach to new or existing retention schedules

hundreds, or even thousands, of categories when classifying content for retention and information lifecycle purposes.

Another approach has as few as five buckets (for example, one bucket each for one year, three years, five years, 10 years, 25 years, and indefinite). The approach is being promulgated by various technology vendors, especially information lifecycle management (ILM) system vendors. ILM is a process for managing information through its lifecycle in a manner that optimizes storage and access by storing content on the most cost-effective media.

John Isaza, Esq., of the Howett Isaza Law Group LLP, observed in 2008 that:

*This 'huge bucket' approach may be useful in managing data storage costs; however, it is not recommended for retention management because it does not effectively address event-driven conditional retention periods and is not granular enough to adequately address federal and state requirements for record retention.*

### **The Optimum Number of Retention Buckets**

Records management and IT professionals are coalescing around 100 record series or buckets as a reasonable number of retention categories for most industries and organizations. Evidence for this conclusion includes:

- In *Records Retention Procedures*, Donald Skupsky, president of Information Requirements Clearinghouse, stated that even a large organization's retention schedule will contain fewer than 100 record series, and 50 to 70 may be sufficient for most organizations.
- Bill Westcoat, who is a senior architect

and records manager subject matter expert for EMC Documentum, has been involved in 60 implementations of records management and electronic document management technology over a five-year period. "I would say 95 percent of them have less than 150 retention categories," Westcoat said. "One client was successful in

going from 7,000 record series down to 300 by leveraging the features and functionality now available in the latest version of their electronic records management software application."

- A 2007 white paper by this author and Lori Ashley, "Streamlining Retention Schedules: The Benefits

# How to Apply the Big Bucket Approach

## Developing New Retention Schedules

1. Identify the physical and electronic records across the enterprise.
2. Organize the records by business function/process into a preliminary classification scheme composed of record series, or "buckets."
3. Conduct legal research to identify relevant federal and state laws and regulations affecting retention and organize them into legal groups that are related by industry and subjects. Approximately 15 to 20 percent of the records series in a schedule will not have legal or regulatory retention requirements, and business requirements for retention will need to be defined. The process of mapping the record series related to the same or similar business processes to their relevant legal, regulatory, and business requirements produces the first draft of the retention schedule.

For a more detailed description of the retention schedule development process, refer to Tina Torres' article on creating process-focused retention schedules in the October 2006 issue of *The Information Management Journal* and Ellie Myler's article on the functional approach for retention schedules in the January 2008 issue.

4. Consolidate the buckets in the draft retention schedule into fewer bigger buckets by assigning the longest retention period among a group of consolidated buckets to its new bigger bucket.

Consolidation into bigger buckets is a collaborative process that requires participation from the legal, records management, IT, and key user departments and requires consideration of many factors including an organization's risk tolerance, history of litigation, and resource constraints.

## Updating Existing Retention Schedules

1. Make sure legal research and business requirements for retention are up-to-date. Refreshing the legal research is recommended every 18-24 months in highly regulated industries.
2. Consolidate record series related to the same and similar business processes and supported by the same or similar legal and business requirements into bigger buckets and assign the longest retention period.
3. Provide crosswalks (tables indicating the relationship between two systems or data structures) to legacy content classified to the old schedule. Legacy content may need to be re-classified to the new retention schedule because any destruction must proceed under the retention schedule and laws existing at the time records are destroyed with a couple of exceptions. According to Donald Skupsky, J.D., president of Information Requirements Clearinghouse:

*In the case where the old retention schedule periods have been used for years to establish destruction dates, the old system cannot be updated to the new retention scheme (e.g., the boxes just reflect the destruction date, not the retention rule or period), and the old retention is generally longer than the new retention, then the organization might elect to 'grandfather' the old retention scheme for these old records only.*

*Alternatively, when the old records are not linked to a pre-existing records retention schedule or the content is generally unknown, and re-classification would be impossible for practical reasons, the organization may elect to destroy the records after some standard, relatively long period such as 10 years after the last activity. This approach lets the 'passage of time' determine the 'valuelessness' of the records for any purpose without incurring the expense of re-classification to the new retention schedule.*

of Big Buckets,” reported that Cargill Inc., which has more than 158,000 employees, simplified its retention rules by reducing its retention schedule to 10 functional categories and 120 record series.

- The 2007 electronic records survey co-sponsored by Cohasset Associates, ARMA International, and AIIM

included questions about the number of record series and their role in electronic records management. In response to the question about how many different record series the respondents’ current retention schedule had, nearly one-third said they had fewer than 100 record series/retention categories, and

more than half (57%) said they had fewer than 250. Cohasset believes that these findings may be somewhat inflated because 40 percent of the respondents were with organizations with fewer than 1,000 employees and 67 percent were with firms with fewer than 5,000 employees. Nonetheless, there is at least a perception that fewer record series are better than more.

## to Retention Schedules

### Big Bucket Approach Recommendations

*Involve the end user community to make sure consolidations make sense.* For example, accounts payable and accounts receivable are usually combined because they are related to the same business process and retention is mandated to the same laws and regulations. But some companies may keep them separate, putting accounts payable in the supply chain management function/process and accounts receivable in the accounting function/process.

*Don’t co-mingle buckets with fixed chronological retention periods (such as “six years”) and buckets with conditional or event-driven retention periods (such as “as long as in force plus-six years”) in a single bucket.* This is not recommended because it is not practical to apply the retention period consistently, a records management and disposition best practice.

*Consider record volumes, especially with physical records.* In the Nevada gaming industry, for example, retention periods for gaming records are mandated by the State of Nevada Gaming Control Board, and most records have a retention period of five years. It may appear reasonable to have a single record series for gaming records with a five-year retention period.

There is, however, one content type, slot machine payout tickets (wagering instruments), whose mandated retention period is only 90 days. The volume of tickets grows rapidly, and it would consume expensive casino real estate or offsite storage if the tickets were consolidated with other gaming records into one bucket and retained for an extra four years and 266 days. In this case, the recommended approach is to create a bucket just for the payout tickets with a 90-day retention period and another for the rest of the gaming records with a five-year retention period.



### Benefits of Bigger Retention Buckets

The potential benefits of big bucket retention schedules are compelling:

*They make it easier for users to classify content for retention.* An abundance of choices does not necessarily lead to better decisions, whether in classifying e-mails, selecting a pair of shoes, or buying pickles. Too many choices can lead to a random selection, a default (but sometimes incorrect) selection, or a refusal to make a selection. Offering users fewer, better-defined choices for classifying content is more likely to lead to accurate and consistent selections, less frustrated and more confident users, and simplified training.

*They make it easier for technology tools to classify content for retention.* With automatic categorization tools and automatic classification engines, ambiguity is reduced and accuracy is increased when these technology tools are presented with fewer, better-defined choices.

*They lower the total cost of ownership.* Less complexity in a retention schedule requires less training and takes less time for periodic maintenance, which lowers the total cost for an organization to maintain a retention schedule.

*They mitigate the risk of retaining records too long.* When users are confident they know how to classify content correctly, they are more likely to do it consistently and retain less unnecessary information.

*Approvers are more likely to approve dispositions.* In organizations that require pre-approval to dispose of information,

approvers are more likely to approve destruction because they have more confidence that users are classifying content accurately and consistently.

*They make it easier to apply retention, especially in enterprise resource planning (ERP) systems such as SAP and Oracle since they can accommodate conditional retention periods.* If a piece of equipment is a business object in an ERP, the retention period is “life of the equipment plus five years,” and the trigger event is when the equipment is sold or disposed, so the ERP can pull the trigger for final disposition.

### Reservations About Big Retention Buckets

The most common argument against big buckets in retention schedules is that information within an aggregated record series or bucket may be retained longer than legally required. Retaining information longer than necessary can leave more records available for electronic discovery and increase storage costs for physical records. For example, Dwight Wallis, a Portland, Ore., records administrator, made the following argument:

*...the disadvantage in more complex environments is that this type of approach can result in the unnecessary retention of large amounts of records, as retention “big buckets” default to the longest retention requirement of their components.*

The counter argument, however, is that more granular retention schedules discourage classification and leave a large volume of unclassified content available for review in court-ordered legal discovery. Christian Meinke, CRM, with Southern California Edison Business Resources, said:

*...if the big bucket approach allows you to index (and presumably better manage) these records, some of that risk might be mitigated despite the volume*

*of records being large, i.e. the process allows me to identify and isolate potentially relevant documents better so discovery costs are reduced.*

Another argument against big bucket record series is that organizations will be unable to locate, retrieve, or dispose of records when needed if the retention schedule is less granular. This may have been true with physical

**The most common argument against big buckets in retention schedules is that information within an aggregated record series or bucket may be retained longer than legally required.**



records management when organizations relied heavily on the descriptive information embedded in retention schedules for locating content. With enterprise content and records management systems, organizations can leverage additional metadata (such as creation date, creator, role, business function, or server) and powerful search tools (such as Google and Autonomy) so descriptive information embedded in the schedule is not the only metadata available for searching.

A final argument against bigger retention buckets is the changing legal environment. That is, if the legal requirements for a record series change, there is concern that it may not be possible to “undo” the bucket and apply the new requirements to the right records. A posting from the records management listserv from Laurie Carpenter, CRM, who at the time was a compliance manager with Koch Industries, colorfully described the potential problem:

*If you take a bucket of yellow paint and a bucket of blue paint, it only takes a few seconds to mix them together into one green bucket. How long and how much would it cost to separate it back out?*

Using the approach described in the sidebar article in which records are organized by business function/process into a classification scheme comprised of record series or buckets, the entire record series or bucket would inherit the changes in a retention requirement. If it made sense to “undo” the bucket, either due to the retention length or volume of the records, it can be done by creating a new bucket. Legacy content classified to obsolete record series will also need to be adjusted. Changing legal, regulatory, and business requirements for retention is to be expected, and a process for implementing changes will be necessary regardless of whether the schedule has 100 or 1,000 record series.

In the final analysis, organizations

need to weigh the odds of end users properly classifying content against risks of potentially retaining content longer than necessary. These risks are unique to each organization and are based on its history of litigation and regulatory scrutiny, its culture for risk tolerance, and its resourcing constraints.

### What to Do Now

Increasingly, organizations are acknowledging that all recorded information retained by employees is potentially discoverable in a court-ordered legal discovery or regulatory investigation, and retention must be applied to all information formats and all media throughout the information lifecycle.

Applying retention to “official records” is already a challenge. The complexity of managing all recorded information for retention purposes is exponentially more challenging. Another key change is the availability of a variety of additional metadata captured in enterprise content and records

management systems, which may make location, navigation, and retrieval of information for litigation and other operational needs much more practical than it was in the world of physical records.

To address the challenges of managing all recorded information and leveraging the advantages of enterprise content and records management systems, organizations need to establish a cross-business function team, including representatives from legal, compliance, IT, records management, the user community, and other key stakeholders, if one is not already in place. The team needs to lead the effort to create a unified enterprise content and records management strategy by developing:

- Information infrastructure with a common taxonomy to organize, describe, and link records (classification scheme), standardized indexing (metadata), and a single set of retention policies for all physical and electronic records, opti-

mized for implementation with electronic records (big bucket retention schedule)

- Technology infrastructure including an enterprise content management and records management solution architecture
- Approach for classifying content for retention and retrieval, preferably transparent to users by using techniques such as templates with embedded classification and automatic categorization by role, department, and/or security profile
- Training and change management strategy, including the identification of information coordinators or specialists to become the local experts in rolling out the plans

The team will also need to assess the long-term impact of migration to 2007 versions of Microsoft Office and SharePoint (the fastest growing product in Microsoft history) on records and information management practice. Organiza-

tions are inheriting (at no cost, if they have already licensed the Microsoft Office 2007 suite) additional capabilities traditionally defined as enterprise content management and will need to determine how to leverage the records management functionality.

Several high-profile organizations have launched successful enterprise content and records management programs. The reality for most organizations is that their leaders do not yet have the appetite for launching an organization-wide content and records management initiative. The Gimmel Group's Mike Alsup, writing for OmniRIM Solutions Inc.'s February 2008 newsletter, said:

*Management is not yet ready to spend enough money to train users and provide the time for them to properly identify and organize existing content, including declaring and classifying these documents as records. Advanced search and categorization tools will become more and more critical in 2008 as a means of searching through the chaos of desktop, shared drive, e-mail, and document repositories to classify existing documents.*

### So, Where's the Fight?

The big bucket controversy appears to be about how big "big" is. When "big" is defined as about 100 buckets or record series in a retention schedule, stakeholders generally agree that having fewer retention buckets is better than having more. The challenge of managing electronic records with enterprise content, records management, and collaboration tools is bringing many long-standing records management practices, such as developing retention schedules, to the forefront for interpretation in a manner appropriate to electronic records.

By combining the electronic ways of working and traditional records management practices, bigger retention buckets should result in a system in which users can manage all recorded information with confidence and integrity. From this author's perspective, the fight is over; it's time to move on! ■

*Susan Cisco, Ph.D., CRM, is a director in Gimmel Group's Enterprise Content and Records Management (ECRM) services organization focused on helping large enterprises improve business performance and compliance through the effective planning, implementation and integration of: enterprise content and records management, geographic information systems, web design, and user experience. She brings more than 25 years of experience in the records and information management field as a practitioner, educator, and consultant. Cisco has successfully consulted with companies in multiple industries, including oil and gas, hospitality, healthcare, insurance, utilities, and government. She holds an M.L.S and Ph.D. in Library and Information Science from the University of Texas at Austin. She is a Fellow of ARMA International. She may be contacted at susan.cisco@gimmel.com.*

## References

- Alsup, Mike. "ECM & RM Trends 2008." OmniRIM Solutions Inc. newsletter, Issue 2: February 2008.
- Carpenter, Laurie. "Re: Bucket Theory." Posting to records management listserv, 25 October 2007. Available at [www.lists.ufl.edu/cgi-bin/wa?A2=ind0710d&L=recmgmt-l&T=0&P=26015](http://www.lists.ufl.edu/cgi-bin/wa?A2=ind0710d&L=recmgmt-l&T=0&P=26015).
- Cisco, Susan and Lori Ashley. *Streamlining Retention Schedules: The Benefits of "Big Buckets."* White Paper. Boston: Iron Mountain, 2007.
- Flynn, Chris. "Re: Bucket Theory." Posting to records management listserv, 24 October 2007. Available at <http://lists.ufl.edu/cgi-bin/wa?A2=ind0710d&L=recmgmt-l&P=13040>.
- Gerard, Elliot. "Implementing Records Retention at General Mills." Presentation at Microsoft Office SharePoint Conference, 5 March, 2008.
- IBM Corporation. "Information on Demand: Information Lifecycle Management (ILM)." IBM brochure, July 2005. Available at <ftp://ftp.software.ibm.com/common/ssi/pm/sp/n/tss-01219usen/TSS01219USEN.PDF>.
- Isaza, John, Esq. E-mail communication with author, 28 March, 2008.
- Meinke, Christian. "Re: Big Bucket Retention – Or Toward a Limited Set of Retention Periods." Posting to records management listserv, 1 June 2007. Available at [www.lists.ufl.edu/cgi-bin/wa?A2=ind0706a&L=recmgmt-l&T=0&P=2024](http://www.lists.ufl.edu/cgi-bin/wa?A2=ind0706a&L=recmgmt-l&T=0&P=2024).
- Myler, Ellie. "Minimizing Risks Through a Corporate Information Compliance Initiative." *The Information Management Journal*, Vol. 42, No. 1, January/February 2008.
- National Archives and Records Administration. "Strategic Directions: Flexible Scheduling," January 2004. Available at [www.archives.gov/records-mgmt/initiatives/flexible-scheduling.html](http://www.archives.gov/records-mgmt/initiatives/flexible-scheduling.html).
- Skupsky, Donald S. E-mail communication with author, 25 March, 2008.
- \_\_\_\_\_. *Records Retention Procedures*. Denver: Information Requirements Clearinghouse, 1990.
- Torres, Tina, CRM, PMP. "Creating a Process-Focused Retention Schedule." *The Information Management Journal*, September/October 2006.
- Wallis, Dwight. "Re: Record Function, Record Type, Record Series and Big Bucket." Posting to records management listserv, 28 September 2007. Available at <http://lists.ufl.edu/cgi-bin/wa?A2=ind0709D&L=RECMGMT-L&P=R7286&I=-3>.
- Westcoat, Bill. E-mail communication with author, 31 January, 2008.
- Williams, Robert F. and Lori J. Ashley. *Electronic Records Management Survey; Call for Collaboration*. Chicago: Cohasset Associates, 2007. Available at [www.merresource.com/pdf/survey2007.pdf](http://www.merresource.com/pdf/survey2007.pdf).