



# Seven Steps to Designating Owners of Unstructured Data

Written by Randy Franklin Smith, president and CEO of Monterey Technology Group, Inc., and Microsoft MVP

Many organizations are seeing surges in the amount of unstructured data in their environments, even as new data breaches come to light every week. As a result, those organizations face increased audit and regulatory pressure regarding loose access controls over unstructured data that might contain sensitive information such as Social Security numbers, credit card data, health care information and proprietary data.

Without a designated data owner, all data must be treated as

mission-critical. That means extra costs in the form of expensive storage, backups and so on. But more important, you can't be secure or compliant without designated data owners. After all, only information owners can decide who should have access to unstructured data. But what if you haven't designated any data owners? Who should they be?

This paper will take you through a seven-step process for establishing information owners for unstructured data, as outlined in Figure 1.

Throughout, we'll show how One Identity Manager - Data Governance Edition helps automate the complex processes of governing and controlling unstructured data.

## Step 1. Identify Unstructured Data Stores

Before you can secure data, you need to know where it is. The notion of location is particularly difficult with unstructured data because most of it is created by end users, with no direct involvement or planning by IT.

In most file-sharing environments, end users might be unable to publish new shared folders. However, they can and do create new subdirectories, which they use to store and share new types of information that ideally should be published as discrete resources, each with its own data owner and access controls. Likewise, Microsoft SharePoint allows end users to create and publish new unstructured data stores independently, through self-service site creation.

Therefore, the first step in getting control of unstructured data is to catalog all the unstructured data stores throughout your network. This paper focuses on the two most common ways that organizations store unstructured data:

- Shared folders
- SharePoint document libraries

## Cataloging shared folders

Shared folders are a computer-level resource, so you need to start by compiling a list of all computers in the environment and their shared folders. Windows offers no built-in report to provide this information, so you might want to turn to PowerShell for this capability.

First, you need a list of computers, and the likely place to start is Active Directory (AD). You might use a script (such as the Get-Shares script) that enumerates all computer objects in an AD organizational unit (OU) and then queries each computer for a list of its shares. For each share, you need the following information:

- Computer name and domain
- Share name
- Share description
- Underlying local file system path

Keep in mind that there might be additional domains on your network or even stand-alone servers with shared folders. And make sure to include any file-sharing appliances, such as NetApp or EMC devices.

## Cataloging SharePoint document libraries

Documenting SharePoint document libraries is a complicated process. First, you must “walk” the SharePoint object hierarchy to find all document libraries. SharePoint’s structural hierarchy from top down consists of these items:

- Farms
- Site collections
- Sites (also known as Webs)
- Subsites
- Document libraries

Again, PowerShell is a tool you can use for recursing through site collections and subsites looking



Figure 1. The seven-step process for establishing owners of unstructured data

for each document library. (The scripts in this article provide an example.) For each document library, you need at least this information:

- Farm
- Site collection name
- Document library URL

## How One Identity can help

Identity Manager - Data Governance Edition scans your file servers, file-sharing appliances and SharePoint site collections to automatically identify unstructured data stores.

### Step 2. Analyze Potential Owners

At this point, you have a list of all shared folders and document libraries, so you've documented all the most likely places where unstructured data might reside on your network. The next step is to analyze these data stores to determine likely owners.

The best owner for a given store of unstructured data is someone who understands the information and works with it regularly (or whose direct reports work with it). The owner needs to be at an organizational level with the authority to make entitlement decisions, as well as the perspective to take into account the business and security implications of granting access to this information.

To find this person, you need to analyze the unstructured data store and its metadata, essentially asking four questions:

- What type of information is in the store?
- Who can access the data?
- Who regularly accesses the data?
- Is the data subject to information security policies?

## What type of information is in this store?


Determine the dominant file types within the document library or folder. For shared folders, WinDirStat is a useful open-source tool for graphically rendering folder structure, data size and file types. You can quickly see how much data is present, how it's organized into folders, and which file types are represented.

However, if all the documents are a generic format (such as Microsoft Word or PDF), you'll need to dig deeper by actually looking at the contents of the data. Your goals are to determine the most important types of stored documents, to understand their business importance to the organization, and to find out what sensitive data (if any) resides in them. You might need to interview people who are frequent users of the data, which brings us to the next question...

## Who has permission to access this data?

Obviously, anyone who uses a given data store must have permissions to it before he or she can access it. So, to identify people who should know more about the unstructured data within a given store, look at the permissions on the folder or library in question. Ideally, you'll document current permissions for each data store.

Typically, the access control list (ACL) for folders and libraries will list one or more groups, each with specific entitlements. Your next step is to understand the membership of each group. Be aware that on Windows file servers, permissions may be granted to local groups unique to that computer. A better practice is to use AD domain groups.



The **best** owner for a given store of unstructured data is someone who understands the information and works with it regularly (or whose direct reports work with it).





To find out who actually uses a document, you can use access auditing.

Likewise, SharePoint supports both AD domain groups and SharePoint groups unique to SharePoint.

Just because a given group has access to a data store doesn't mean that all its members access the information. Entitlements are commonly much broader than necessary. This happens because of the absence of a knowledgeable data owner, because busy administrators sometimes lack an understanding of the data and business requirements, and simply because permissions become outdated over time. Therefore, a data store's permissions might not help you zero in on the key users of that data. However, documenting the current entitlements on the data store is still a necessary step, as you'll see later in the process. But the next question provides an effective way to find the real users of a given data store.

### **Who regularly accesses this data?**

To find out who actually uses a document, you can use access auditing. Both Windows NTFS and SharePoint provide an audit capability. By enabling auditing for a period, you can analyze the logs for usernames that show up frequently. Enabling auditing on either platform requires access to the administrative controls and is complicated.

**NTFS auditing** — Windows provides two audit categories for auditing access to shared folders. You can enable the Detailed File Share subcategory on a given system, and Windows will begin recording every file access for all shared folders on that computer with event ID 5145, which logs the username, computer name, shared folder and file name. Event ID 5145 also logs the type of access (such as Read,

Write or Delete) so that you can distinguish between users who produce and modify data, as opposed to those who just read it.

For more granular control over what activity is audited, you can use the File System audit subcategory. If you use this subcategory, you'll need to define audit policy on each folder you want to track, specifying who to audit and which types of access to track. This category produces event ID 4663, which logs essentially the same information as ID 5145.

File auditing in the Windows security log is complex because you must either cope with the system auditing every access to every file (Detailed File Share category) or configure audit policy on each folder (File System category). Either way, the events logged by Windows are famously cryptic and have a high degree of noise and duplication. Furthermore, each computer records security events to its local log. Ultimately, there's no way to effectively analyze access events without knowledge of the arcane Windows security log. Plus you need a log management tool to consolidate logs from multiple systems and perform the filtering and summarization necessary to identify the key users of a given folder.

### **SharePoint auditing —**

SharePoint auditing is controlled at the site collection level. SharePoint farms often have thousands of site collections, and enabling auditing is a manual operation accessed through each site collection's Site Collection Administration pages. With SharePoint, you choose which types of access (such as View, Create, Update or Delete) to audit for the entire site collection. The audit process includes all objects, which means — particularly in

the case of View access — that SharePoint records multiple events in its internal audit log, not just for documents, but for list items and every other page view.

As you can see, Windows and SharePoint auditing provides the information you need to identify key users of a given unstructured data store, but both auditing systems require skill and resources. After you've identified these users, interview them about the information and how they use it. Determine whether your conclusions about the data and its sensitivity are correct. Analyze the users' departments and job titles to see whether they're at an appropriate level to be the data owners, or whether you should look to their managers.

Is this data subject to any existing information security policies or classifications? Once you have an understanding of the kind of information a data store contains and a list of the primary users of the data, it's time to determine whether this information fits into a previously identified type in your organization's information classification scheme. You also have to consider whether it's subject to any regulatory compliance requirements.

By comparing what you know about the data store with your organization's information security policy, classification and compliance documents, you should be able to determine whether this data store should be singled out for special treatment as defined by existing policies. Such policies might mandate who should make entitlement decisions, thereby simplifying the process of ownership selection. At the very least, you'll want to list any applicable policies, regulatory requirements and classifications

in your documentation for this data store.

## How One Identity can help

Identity Manager - Data Governance Edition automates the entire tedious process of owner analysis. There's no need to wrestle with Windows and SharePoint auditing or to manually collect and document permissions. The software automatically collects all this information — and then goes further by analyzing it and automatically suggesting the best owner.

Further enhancing the solution is the Classification Module for Identity Manager - Data Governance Edition, which automatically identifies and protects sensitive data based on policy, eliminating the need to open and examine each file manually. Its MRI-like ability to find, identify and classify unstructured content manages risk, enforces security and takes the stress out of audits.

### Step 3. Confirm Ownership

After you answer the questions in Step 2 about a given data store, you should have one or more prospective data owners in mind. Now it's time to get formal acceptance of ownership, which might be as simple as getting the owner to respond affirmatively to an email message.

But, depending on the sensitivity of the data, the position of the prospective owner, and any applicable information security policies, formal ownership confirmation might require more. For example, you might need approval from executives in concerned business areas, confirmation by the information security officer or approval from the compliance officer.

Throughout this entire process, you've been building a list of unstructured data stores. Each data store on the list should include:

- Name and description
- Network location
- Current permissions
- Relevant compliance and information classification endorsements
- Designated owner (along with the ownership confirmation and date)
- Last attestation

## How One Identity can help

Identity Manager - Data Governance Edition eliminates the email-heavy burden of confirming ownership, thanks to its built-in workflow engine. After an owner is selected, the software sends the automated message, "We've identified you as a data owner; is this your data?" If the owner fails to respond within a certain time, the product can automatically escalate the workflow by contacting the owner's manager or taking other action.

### Step 4. Begin Initial Attestation

The first act of the confirmed owner of a given unstructured data store should be an initial attestation to current entitlements or permissions that are in place. You'll need to provide the data owner with a list of all assigned data stores, along with the information you collected earlier — most importantly, a report of the current permissions.

The data owner might need additional information, such as the membership of groups referenced in the permissions. Obtaining this information and providing it in a readable format to the data owner isn't complicated, but can be time-consuming. This is especially true with local groups on file

servers and in SharePoint, or when dealing with nested groups whose members include other groups.

Now, the owner should approve or reject each entitlement and specify any other permissions that should be added. Save this attestation for use in the next step, but also as permanent documentation for future audits and compliance reviews.

### How One Identity can help

Identity Manager - Data Governance Edition automatically prompts a new owner to review current entitlements on the unstructured data store. The owner can accept or even delegate approval to someone else, in effect saying, "I'm not the owner of this, but I know who is. It's John. Go ask him."

There's no need for the administrator to document permissions and group memberships and then pursue the data owner to complete attestation. The software manages the entire process and provides an easy-to-use interface to the non-technical data owner with all the information necessary to make good business decisions.

### Step 5. Remediate Access

This step should simply be a matter of adjusting data store permissions to reflect the changes that the data owner specifies. However, it is yet another step in the long process between IT and the data owner, and another area where the process can stall.

After the changes are made, the attestation process should be repeated until the owner confirms that all current entitlements are appropriate. Naturally, that attestation in particular is important to save for future audit and compliance reviews.

### How One Identity can help

Identity Manager - Data Governance Edition completely eliminates Step 5. When the owner specifies entitlement changes in the previous step, the software adjusts permissions and group memberships automatically and immediately.

### Step 6. Implement an Owner Approval Process


When you reach this step, you've already achieved an important milestone in securing unstructured data. However, initial attestation and remediation is only a point in time — requests for new access are constant. Therefore, it's crucial that you implement an owner-approval process going forward for all changes in the data store's access controls.

Without an enterprise identity and access management (IAM) solution, this is typically a laborious process that depends on email and strict observance of manual steps. Some organizations without a real IAM system create an access request list in SharePoint and set up workflows to help prompt the requesting party, owner and administrator through the various steps in the process.

However you implement the owner approval process, you need a business record for each entitlement change that shows who requested the entitlement, the trustee to be assigned access, approval by the owner and execution by the administrator. This record is necessary so that at any time you can justify why a user had access to a given data store, or for a given date, who had access and why.

### How One Identity can help

Identity Manager - Data Governance Edition's IT Shop



Requests for new access are constant. Therefore, it's **crucial** that you implement an owner-approval process going forward for all changes in the data store's access controls.



uses an e-store metaphor to simplify and automate the process of quickly fulfilling access requests while ensuring data owner approval and relieving IT from unneeded involvement. The user simply visits the IT Shop and requests access to a new resource. The software contacts the data owner with all necessary information and carries out any needed entitlement changes.

## Step 7. Implement an Owner Approval Process

Even with direct owner involvement in access requests, permissions tend to drift away from current business requirements because of events such as changes to business processes, department reorganizations and job changes.

Enterprise IAM systems such as Identity Manager - Data Governance Edition can catch many of these events and automatically adjust entitlements or prompt data owners to review them based on new criteria. If you lack an enterprise IAM solution, the only way to keep up with these changes is to have the data owner periodically review current entitlements and specify any needed changes in the light of current requirements. In fact, even with an enterprise IAM tool, many organizations require regular re-attestation for critical

information security resources to ensure compliance.

Re-attestation is usually conducted at least once per year, but might be required on a quarterly basis. The process is similar to that described in Steps 4 and 5. The difference is that a schedule is needed to trigger the re-attestation process, and you must follow up with owners who fail to respond in a timely manner. Collecting the information for each re-attestation can be time-consuming.

## How One Identity can help

Identity Manager - Data Governance Edition supports traditional periodic re-attestation, but it can also replace re-attestation with something far more efficient and secure: event-based access management. Between re-attestation periods, the software detects job changes and other business events, and automatically remediates access to the system manager or appropriate data owner.

## Identity Manager - Data Governance Edition

Identity Manager - Data Governance Edition protects your organization by giving access control to the business owners who know who should have access to which sensitive data. Business owners are granted the power to analyze, approve and

fulfill unstructured data access requests to files, folders and shares across NTFS, NAS devices and SharePoint.

Identity Manager - Data Governance Edition helps data owners (not IT) determine who should have access and automates the request-and-approval workflow, keeping your organization from being the next security headline — all while reducing the burden on IT. Some key capabilities of the solution include:

- **Restricted access** — Define access policies for your organization to ensure that sensitive unstructured data is accessible only to approved users. Identity Manager - Data Governance Edition locks down sensitive data such as files, folders and shares across NTFS, NAS devices and SharePoint.
- **Data owner assignment** — Determine and assign the appropriate owner of data for all future access requests by evaluating usage patterns and read and write access.
- **Simplified auditing** — Identify user access to enterprise resources such as files, folders and shares across NTFS, NAS devices and SharePoint to provide key information during audit preparations.
- **Automated access requests** — Use built-in workflows to automatically direct access requests from the request portal to the appropriate data owner. Approved requests are automatically and correctly fulfilled, with no burden on IT.

Display name	Object type	Risk Index
\\THEWHPHLE\CS\Tee Set 10000\Social Insurance Number 118-123-321 (1).txt	Governed Data	0.35
\\msagrd-01n-hat.ca.spff\Users\Administrator	User account	0.25
Admin, Secondary	Employee	0.20
\\msagrd-01n-hat.ca.spff\Users\theladnld	User account	0.20
\\msagrd-01n-hat.ca.spff\Automation Users\Autonomous	User account	0.20
\\msagrd-01n-hat.ca.spff\Automation Users\Q1MAdmin	User account	0.20
\\msagrd-01n-hat.ca.spff\Development\Leigh LV, Vietnam	User account	0.20
\\msagrd-01n-hat.ca.spff\Print Test Users\DomainAdminUser	User account	0.20
\\msagrd-01n-hat.ca.spff\Users\Bathandnld	User account	0.20
\\msagrd-01n-hat.ca.spff\Users\Secondary Admin	User account	0.20

Governed data name	Owner (Employee)	Count activities	Risk Index (calculated)
\\THEWHPHLE\CS\Tee Set 10000\Social Insurance Number 118-123-321 (1).txt		0	0.35

Figure 2. Identity Manager - Data Governance Edition protects your organization by giving access control to the business owners who know who should have access to which sensitive data.

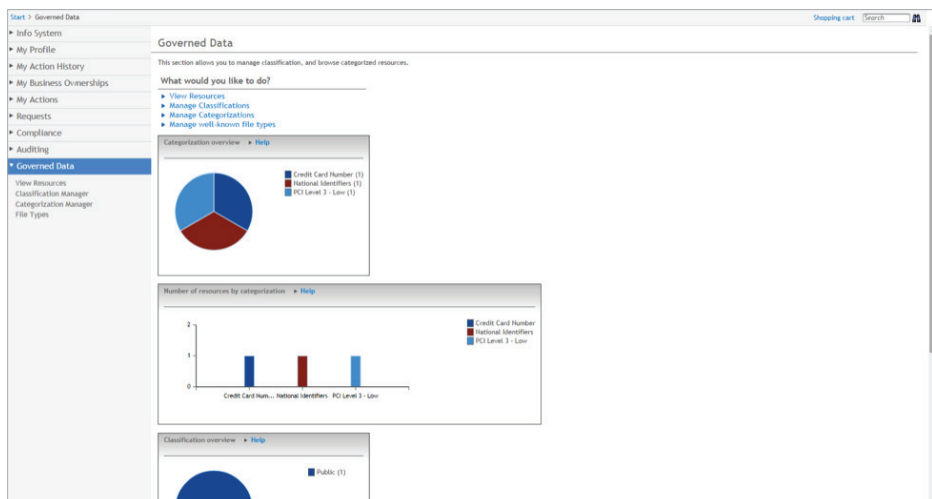


Figure 3. Classification Module for Identity Manager - Data Governance Edition automates the process of securing and classifying unstructured data.

**Access verification**—Ensure that only approved users have access to specific resources, including those who have left the organization or department, or whose roles have changed. Identity Manager - Data Governance Edition enables you to monitor user and resource activity, and configure and schedule a recertification process for data owners to verify and attest to employee access.

**Personalized dashboard**—View trends, historic and current data access activity, and attestation status on a personalized dashboard with reports that you can use to prove compliance to auditors.

## Conclusion

As the volume of unstructured data in your environment grows, it's increasingly critical to protect that data appropriately, as well as to ensure regulatory compliance. To achieve those goals, you need to designate owners who are in the right position to decide who should have access to the data.

Using the seven-step process laid out in this paper, you can identify your unstructured data stores, establish owners for the data, and

ensure that the right users have access — now and into the future. While this process can be difficult and time-consuming with manual approaches and native tools, it can be streamlined and automated with comprehensive IAM solutions like Identity Manager - Data Governance Edition.

## About the author

Randy Franklin Smith is an internationally recognized expert on the security and control of Windows and Active Directory security. Randy publishes [www.UltimateWindowsSecurity.com](http://www.UltimateWindowsSecurity.com) and wrote "The Windows Server 2008 Security Log Revealed" — the only book devoted to the Windows security log. Randy is the creator of LOGbinder software, which makes cryptic application logs understandable and available to log-management and SIEM solutions.

As a Certified Information Systems Auditor, Randy performs security reviews for clients ranging from small, privately held firms to Fortune 500 companies, both national and international. Randy is also a Microsoft Security Most Valuable Professional (MVP).



## For More Information

© 2017 Quest Software Inc.  
ALL RIGHTS RESERVED.

This guide contains proprietary information protected by copyright. The software described in this guide is furnished under a software license or nondisclosure agreement. This software may be used or copied only in accordance with the terms of the applicable agreement. No part of this guide may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording for any purpose other than the purchaser's personal use without the written permission of Quest Software Inc.

The information in this document is provided in connection with Quest Software products. No license, express or implied, by estoppel or otherwise, to any intellectual property right is granted by this document or in connection with the sale of Quest Software products. EXCEPT AS SET FORTH IN THE TERMS AND CONDITIONS

AS SPECIFIED IN THE LICENSE AGREEMENT FOR THIS PRODUCT, QUEST SOFTWARE ASSUMES NO LIABILITY WHATSOEVER AND DISCLAIMS ANY EXPRESS, IMPLIED OR STATUTORY WARRANTY RELATING TO ITS PRODUCTS INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NON-INFRINGEMENT. IN NO EVENT SHALL QUEST SOFTWARE BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL, PUNITIVE, SPECIAL OR INCIDENTAL DAMAGES (INCLUDING, WITHOUT LIMITATION, DAMAGES FOR LOSS OF PROFITS, BUSINESS INTERRUPTION OR LOSS OF INFORMATION) ARISING OUT OF THE USE OR INABILITY TO USE THIS DOCUMENT, EVEN IF QUEST SOFTWARE HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Quest Software makes no representations or warranties with respect to the accuracy or completeness of the contents of this document and reserves the right to make changes to specifications and product descriptions at any time without notice. Quest Software does not make any commitment to update the information contained in this document.

## About One Identity

The One Identity family of identity and access management (IAM) solutions, offers IAM for the real world including business-centric, modular and integrated, and future-ready solutions for identity governance, access management, and privileged management.

If you have any questions regarding your potential use of this material, contact:

### Quest Software Inc.

Attn: LEGAL Dept  
4 Polaris Way  
Aliso Viejo, CA 92656

Refer to our Web site ([www.oneidentity.com](http://www.oneidentity.com)) for regional and international office information.