



---

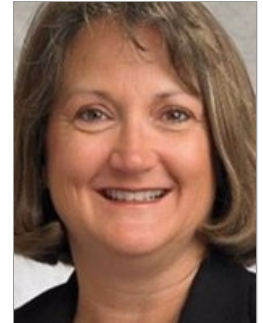
**Portfolio Media, Inc.** | 111 West 19th Street, 5th floor | New York, NY 10011 | [www.law360.com](http://www.law360.com)  
Phone: +1 646 783 7100 | Fax: +1 646 783 7161 | [customerservice@law360.com](mailto:customerservice@law360.com)

---

## How To Jump-Start Your Info Governance Initiative

By **Linda Sharp, ZL Technologies**

Law360, New York (January 9, 2017, 11:25 PM EST) -- In a time of increasing budget cuts, how can we possibly implement an information governance initiative and address the looming EU General Data Protection Regulation requirements? One of the most critical factors in tackling both of these costly initiatives is an understanding of what data you have and where it is maintained. At first blush, organizations might look to their structured data stores, but generally organizations have a pretty good grasp of what is maintained in these environments. It is the unstructured data stores — emails, file shares and SharePoint sites — that are ripe for reaping havoc when it comes to meeting GDPR or implementing an information governance initiative.



Linda Sharp

Why? Because as a general rule, you have no idea what in the world is actually stored in them. Thus, you cannot make a viable determination about what needs to be cleaned out, what might have a drastic effect when it comes to e-discovery, or what data may place you at risk under GDPR and other regulatory requirements.

### Looking Forward to GDPR

So with all this talk about GDPR, what does it actually mean for most companies? In short, among other requirements, GDPR will soon require companies with EU citizen personally identifiable data to accommodate a new set of data subject rights, including:

- A subject's right upon request to receive any personal data of theirs being processed
- A subject's right to have their data deleted
- A subject's right to have their data transferred
- The deletion of personal data once it is no longer being used for its original purpose

Though meeting these requirements might seem to be a daunting task, the good news is that implementing an information governance initiative will go a long way toward reaching GDPR compliance. You will soon see how these two go hand in hand.

### Time for Spring Cleaning

Today, I'm focusing on those ever-growing file shares. Think back to when you first started at your organization. You were handed your security badge, shown around the facility, given your email address, and perhaps quickly pushed through some form of training on company policy, etc. But you were also handed a link to your own personal file share. As is commonplace in most organizations, this process was replete of any training, any understanding as to how this data might be controlled by your "system of record," if you even had one. You were told that it is your file share where you can "put your stuff."

It is not uncommon to hear tales from organizations that complain of file shares in the hundreds of terabytes and even petabytes of data, all of which no one has any idea what is contained therein. In organizations with any longevity, approximately 50 percent of the data being stored is from

employees that have long since departed the organization. This is compounded by the knowledge that such a tremendous percentage of this data has passed its useful life and is infiltrated by data that is employee personal “junk.”

This problem is compounded as departments and working groups create department- or project-specific file shares. No one person actually owns responsibility for these data stores, and if someone did at any point, that person has moved on.

As we look to ensure that our companies are meeting GDPR requirements before the drop dead date of May 25, 2018, how can we truly confirm that we are ready and will be compliant if we have no idea what sits in these enormous data stores? Furthermore, how can we implement a viable info gov initiative with this black hole of data looming over our heads?

So, what’s the solution? Clean it up!

### **Option One: Do-it-Yourself Employee File Cleanup**

I have spoken with organizations that are attempting to clean up these data stores, but yielding blood, sweat and often tears as they have employees go back through and try to identify what they should get rid of and what needs to be kept. It’s their data, so have them do it, right? Though at first blush that seems like a reasonable request, there are several problems with this strategy:

1. How many employees do you have sitting around doing nothing? Can they now spend weeks or months on end trying to clean up their data?
2. Who is going to clean up the data for former employees?
3. What about department- and project-specific data?
4. What about the cost associated with having individuals now spend their working day cleaning up data instead of doing their regular jobs?
5. What, if any of it, is currently subject to litigation hold? Are you really going to rely on an employee to make this determination?

To address the cost concerns, let’s look at a hypothetical mid-sized company that wants to perform file cleanup using the employee do-it-yourself method. Assuming:

- a. 10,000 employees
- b. 80 hours of clean up time (conservatively)
- c. Blended labor rate of \$40 per hour (including benefits, associated taxes, etc.)

In this scenario, the resulting costs of employee self-cleanup would be \$32,000,000. Yup, that’s \$32 million U.S. dollars. You’re still not done yet, though. So now you have the personal shares for the existing employees cleaned up, but what are the costs associated with now cleaning up the department, project and former employees’ data? Who is going to take on that responsibility? Most importantly, are you going to revisit this process again a year or two from now and indefinitely into the future?

### **The Real Solution? Ongoing File Management**

There is a better way. Perhaps now might be the time to truly fix this problem and ensure that your data is appropriately managed on a go-forward basis. Though there are various applications on the market that allow you to do varying passes and depth of analysis across these data stores, it is essential to ensure that viable processes are implemented on a go-forward basis so that you never have to go through this process again. A continuously organized file share not only optimizes all functions of information governance, but also makes finding data for GDPR subject requests much more feasible.

Regardless of a company’s principle reason for file analysis, such initiatives often begin with a scan of metadata. This scan provides users with a comprehensive overview of variables such as location, file type, author, and creation and modification dates, displayed through intuitive and practicable charts and graphs. Based on these metadata reports, the company can use this

information to make immediate decisions on their existing data. They can choose to locate and remediate any "ROT" (redundant, obsolete, trivial) data, search for and preserve legal hold data, or make other metadata-based decisions. In fact, experience has shown that the majority of insight from file analysis typically comes from metadata-analysis alone.

Sampling provides a more in-depth look, enabling organizations to go beyond metadata and analyze content. This process provides an overview of where various information types such as sensitive information (PII, PHI, PCI), potential records, legal hold data, etc., most likely reside. This methodology allows them to confidently prioritize content-based decisions without having to index the entire file share at once. This is important to GDPR compliance, because it informs companies where this information currently resides such that they can make informed decisions about appropriate remediation or where should look when subject requests for data access or deletion arise.

### **Looking Closer at Content**

Though metadata analysis and advanced sampling have their place, companies looking for truly defensible GDPR compliance will need full content-analysis. Content-analysis enables a fully customizable search for criteria such as file type, keyword and character patterns indicative of personal information, such as health and credit card data. At any and all points throughout this process, users can tag files individually or on a large scale with adjustable classification tags, and then perform policy-based remediation actions, such as deletion, relocation to a secure environment, and assignment of access privileges.

Not only does this help companies enable their legal department, mitigate security risks, and optimize other functions of information governance, it is essential to complying with GDPR. If an individual whose data your company has been using requests that you hand over everything you have on them, how are you going to be able to find it without searching through the content of all enterprise data?

### **Moving Forward**

File cleanup is an essential step in information governance and GDPR compliance initiatives, but without ongoing management, data will quickly accumulate once again and demand repeated cleanup. Fortunately, once your organization has determined its remediation actions, some solutions will allow it to use these actions to form ongoing policies, applied on data as it is created. This gives an organization peace of mind that its file share will remain managed as new data enters their system, and therefore it will not have to regularly repeat the file analysis process.

At the end of the day, by looking into and cleaning up your file shares, you benefit in two areas.

First, this is a great way to jump-start your information governance initiative. You'll yield real, hard-core numbers that aid in justifying that ROI for a future initiative. You'll have cut your data storage costs, reduced risk by eliminating data that is no longer useful to the organization, and reduced the costs associated with e-discovery by eliminating those potential targets.

Second, as the deadline for GDPR rapidly creeps upon us, you'll have identified stores of EU personal data, and made a determination to either delete or create a policy for how and where that data will be stored on an ongoing basis.

The time to get started is now!

---

*Linda Sharp is associate general counsel of ZL Technologies.*

*The opinions expressed are those of the author(s) and do not necessarily reflect the views of the firm, its clients, or Portfolio Media Inc., or any of its or their respective affiliates. This article is for general information purposes and is not intended to be and should not be taken as legal advice.*