SEGMENTATION, EXTRACTION, AND DATA MASKING IN CONTENT CLOUD

Managing Data Protection for Content Compliance

An organization's archived content can provide a lot of value. However, retaining this data incurs a fair amount of cost and, because this archived content often contains personal customer information, it introduces potential risks as well. That's why organizations must have a defined purpose for keeping any data, especially personal customer data.

Potential reasons for retaining archive content include:

SATISFYING REGULATORY COMPLIANCE

Some industries (such as banking and finance) have regulatory bodies that require organziations to retain data for a period of years. Often the main reason organizations in these industries retain content so long is to meet these regulations. However, software solutions such as Systemware Content Cloud help these organizations manage archive content, while deriving additional business value from this content as well.

PROVIDING ARCHIVE CONTENT TO CUSTOMERS

Many organizations have found that providing customers online access to their archived data provides tremendous value, as it can simultaneously modernize work processes while improving operating costs. Recent regulations such as the European Union's GDPR also require organizations to make this archived data available, making this even more important.

DRIVING BUSINESS INTELLIGENCE

Archive data can provide valuable insights into customer trends, demoraphics, and more. To best realize this value, organizations must first have a some understanding of what information might potentially prove valuable and have some method to assign relevant metadata to this content for easier retrieval. Otherwise, organizations end up swamped with data that is impossible to shift through.

Why Data Protection is Important

When organizations choose to retain data, it's important to be aware of how they are handling theircustomers' personal information. Not properly handling personal data or storing beyond necessity introduces several risks:

UNNECESSARY COSTS

organizations who needlessly store content incur unnecessary operational costs both from the technology itself and the employees who must manage it.

FINES & PENALTIES

organizations open themselves up potential regulatory fines and other penalties if they mishandle personal data. Regulations like the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act provide protections that require organizations to delete sensitive personal data to reduce potential privacy risks.

DAMAGE TO BUSINESS

Mishandling personal data can expose organizations to bad press and even potential civil liability—all of which results in a loss of reputation and other financial impacts.

The risks are important to be aware of, because loss of personal data can and does happen. Numerous news reports have come about in recent years where organizations have incurred attacks resulting in the loss of personal data for millions of customers. These events have shown that threats to customer data can come from both within and outside the organization.

External Threats

When we think of threats to an organization, we most often think of data exfiltration, where people outside the organization are obtaining customer data. Two common ways this occurs are:

- When points of weakness within a system are exploited to access improperly exposed data
- When the compromised identity of a real employee is used to access properly secured data

Internal Threats

While we wish it were true that that no one within the organization would threaten customer data, threats from "bad actors" inside the organization do exist, either from employees or contractors. Two common sources of data leaks are:

- Users of test or training databases (such as developers, testers, trainees, etc.)
- Users of analytics databases (analysts, researchers, etc.).

Building a Structured Approach to Data Protection

To properly address these threats, organizations must have a structured approach to protecting customer data. It may be tempting to simply rely on instinct when developing data protection and retention policies, an organized approach will provide a scalable solution that can be distributed across the organization, with simple, automated tools like Systemware Content Cloud available to make the process easier.

Here are some questions to ask when building a data protection plan:

Do you have a legal right to process the data? Just because you can access the data does not mean you have a right to retain and use that information. Some personal data must be deleted immediately, while other information can be retained if the data is suitably protected.

1. **Do you have a legal right to process the data?** Just because you can access the data does not mean you have a right to retain and use that information. Some personal data must be deleted immediately, while other information can be retained if the data is suitably protected.

2. **Do you have a regulatory obligation to retain the data?** If you have a regulatory requirement to keep the data, you should retain it only for as long as required and do so securely. Once the regulatory requirement is up, it's best to delete sensitive portions of the data in a timely way to prevent fines and other undue costs.

3. **Is there value within the data that can be suitably be realized?** If not, simple archiving with basic indexing is enough. If so, better indexing can help you locate relevant data faster across multiple repositories.

4. **Do the users viewing the content need access to the personal information?** If there is not a regulatory requirement to keep data that is otherwise useful, organizations should make sure they are maintaining customer privacy. This should be of primary importance. Even if you do need to retain personal information for compliance, organizations should find a way to isolate personal information from other data.

5. **Is there a suitable method for reducing potential risk?** If you are not going to completely remove the data for regulatory requirements or other reasons, is there a method to prevent loss of personal information (data encryption, group and user-based access policies, data masking, low-availability storage, etc.)

Handling Subject Rights Requests (SRRs)

Under laws like GDPR, customers have the right to access their information ("right of access") or be forgotten ("right of erasure"), amongst others. To fulfill these requests, organizations must be able to quickly locate relevant personal information and either package it up for delivery (for a right of access request) or delete it from the system (for a right of erasure request).

systemware

To handle these Subject Rights Requests (SRRs), you need to be able to perform a few vital functions:

IDENTIFY
Quickly locate and package relevant data across repositories and million-page reports

RETRIEVE & REMOVE
Address the relevant portion of content without compromising other data

AUDIT
Track who has accessed or edited all reports and other documents within the system

SECURE
Protect data from bad actors both at rest and in transit.

MASK
Provide sensitive data to only privileged users without storing content twice.

SIMPLIFY
Make routine tasks easier while preventing unprivileged users from viewing sensitive data

Manage Personal Data with Systemware

Systemware offers several capabilities to simplify data protection management:

IDENTIFY CONTENT WITH INTELLIGENT INDEXING

- Storing intelligent metadata about the document allows you to quickly identify data related to a specific account, day, location, etc.
- Index at the document, page, or line level, providing multiple ways to tag and retrieve the same document.
- Package content together and deliver a single file in minutes for an audit or customer request.

RETRIEVE AND REMOVE WITH SEGMENTATION

- Store the document in small segments that are divided based upon an index. Can be the whole document, a page or group of pages, or even a portion of a page.
- Quickly retrieve and (re-)package and deliver portions of the overall document dynamically in a way that is seamless to the user.
- Delete segments (in the case of a right of erasure request) without affecting the rest of the document.

systemware

AUDIT SYSTEM ACCESS WITH CONFIGURABLE LOGS

- Robust and configurable logging of report access, edits, and more.
- Logs stored within Content Cloud, providing the same detailed search, line-level retrieval, and packaging as other content within the system. Quickly locate relevant access records and package as a single PDF.

SECURE WITH DATA ENCRYPTION AND ACCESS PERMISSIONS

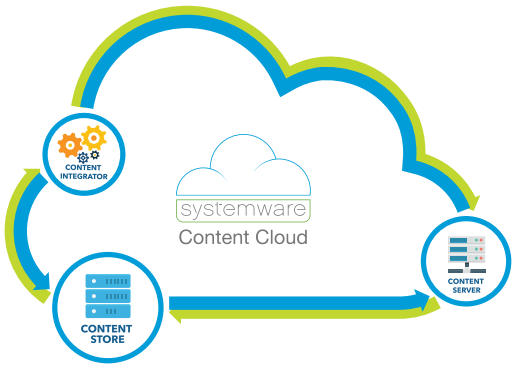
- Encrypt customer data both in transit and at rest.
- Employ user, group, and role-based permissions to control access
- Ensure anyone other than permitted users cannot access the content.

MASK WITH STATIC AND DYNAMIC DATA MASKING

- Dynamic data masking that delivers personal information in a permanent/non-reversible way so that non-privileged users cannot access personal data.
- Multiple levels of masking for different users and access requirements.
- For content not covered by retention requirements but still valuable for use in analytics, organizations can re-archive masked data and delete the raw data, providing anonymization through static data masking.

SIMPLIFY WITH WORKFLOWS AND AUTOMATION

- Using an approach that is essentially "anonymization by design," workflows pass data directly to algorithms (Hadoop cluster, etc.) that process the data automatically and then delete raw source data from data processing server. This prevents exposure of sensitive data to those that do not need to access it.



Systemware Content Cloud lies at the intersection of content services and big data. Archive content from across your organization and manage retention, find and extract data wherever it is stored, then transform and deliver information in the context needed.

Learn more at systemware.com/content-cloud or call to request a demonstration at 844.343.0200