



# Compliance Archiving of Advanced Content Management Systems (CMSs)

THE GUIDE TO CONFIDENTLY CAPTURE PERSONALIZED DIGITAL EXPERIENCES



These days, most people check out a company's website before deciding whether to trust, work with, or buy from them. Prospective customers want to see engaging, user-friendly, dynamic, interactive websites that are personalized to their unique needs and specific to their individual journey. Fortunately, products like [Adobe Experience Manager](#), [Sitecore Experience Manager](#), and the open-source alternative [Drupal](#) give companies a way to provide an experience that draws visitors in and customizes their ongoing interactions with the website, converting them to customers.

These advanced content management systems (CMSs) open the door for marketing teams to be endlessly creative and to refresh their messages as often as they like. As [Sitecore says](#), "a CMS needs to be as powerful as the business it serves."

But if you're in the highly regulated financial and insurance industries, there's something else that needs to be as powerful as your business: your archiving solution. While website technology has enabled new business opportunities, it's also created new risk management challenges. Most archiving solutions haven't caught up.

Here's what you need to know to ensure that your archiving solution keeps pace with your CMS.



# THE ADVANTAGES OF USING AN ADVANCED CMS

An advanced CMS offers a sophisticated, highly customizable approach to website design that can be leveraged by organizations that are serious about using their websites to present their brand, grow their business, and convert visitors to customers. These systems help organizations create extremely interactive, user-friendly, and hyper-personalized experiences for individual customers, informed by each customer's characteristics and history of online interactions with a product, brand, or company. They also track the end-to-end customer journey to determine how users reach the company's site, where they interact, and where they click away so that developers and marketers can create more effective content.

## A MORE DYNAMIC—AND MORE ENGAGING—USER INTERFACE

Interactive content isn't exclusive to websites designed with advanced CMSs, but these systems take dynamic design elements to a new level. In addition to mouse-over and hover-over features that require user input, advanced CMSs may incorporate text, images, video, audio, and other multimedia content. These webpages increasingly function as dynamic web applications rather than static single pages—which ups the complexity of demonstrating compliance.

## PERSONALIZED WEBSITES

In addition to these on-page dynamic features, advanced CMSs enable multiple versions of webpages that can be personalized to the individual visitor. Just like a Spotify feed grows more and more specific to a particular user's preferences and tastes over time, websites designed with advanced CMSs adapt to users as they visit to ensure that they're seeing content that's relevant and helpful to them.



“Consumers feel that digital experiences have fallen short of expectations, yet they’re more likely to shop with a brand that treats them in a personal manner.”

*[Accenture Interactive 2018 Personalization Pulse Check](#)*

The rewards are evident: research from [Accenture](#) shows that “91 percent of consumers are more likely to shop with brands [that] recognize, remember, and provide [them with] relevant offers and recommendations” than with brands that treat them as generic shoppers. [Gartner even predicted](#) that “by 2020, smart personalization engines used to recognize customer intent will enable digital businesses to increase their profits by up to 15 percent.”

How do these features work? Adobe Experience Manager boasts that it [uses artificial intelligence and machine learning](#) to design and manage websites that monitor and adapt to individual users’ precise interactions with the site. Those algorithms make certain assumptions about users based on their global location, their history (if any) with the site, and even their interactions with the site on the particular day in question. This enables dynamic site adaptation with flexible navigation that includes text, images, and video.

Software like Adobe Experience Manager also gives marketers tremendous power and flexibility in understanding how customers are using their website as well as varied customizations based on that improved understanding. This allows organizations to create extremely detailed rules for how users navigate sites and what they see as they do, creating personalized experiences that are customized across the lifetime of the user.

For example, an insurance company that offers different product lines and rates depending on a customer’s state, age, and gender can detect a visitor’s location based on their IP address and immediately offer a customized array of their available options. If a customer provides additional information by interacting with the site, perhaps by selecting from a pull-down menu, the site can then store that data for future use, providing an individualized site.



**Which advanced CMS should you choose?** Gartner has consistently rated Sitecore and Adobe at the top of its Magic Quadrant for [Web Content Management](#), giving Adobe slightly more points for “completeness of vision” and crediting its advances in artificial intelligence and machine learning. However, Sitecore ranks well in its training programs. And Acquia, which offers the open-source Drupal platform, is right up there with Sitecore and Adobe, earning high marks for its personalization and customer journey capabilities. To determine which is best for your organization’s needs, compare reviews on websites like [G2](#), [Capterra](#), [TrustRadius](#), and [Gartner’s Peer Insights](#).

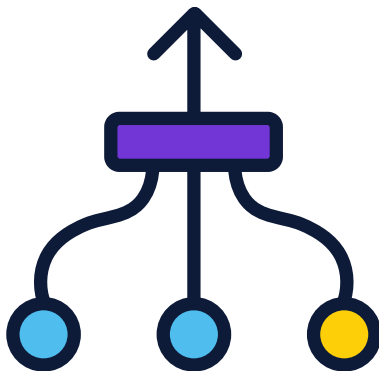
## TESTING AND ANALYTICS

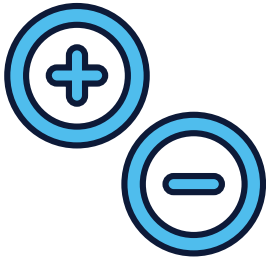
Advanced CMSs allow marketers to offer A/B testing, comparing the conversion performance of different page options and collecting metrics about how customers respond to content. This, along with general user tracking, helps marketers and developers build the best and most compelling website possible, informed by real customer experiences and data.

## STREAMLINED SITE UPDATES

One of the major draws of advanced CMSs is their ability to manage pages and content, streamlining the administrative tasks associated with website updates. This makes it fast and easy for marketing teams to update their content, allowing new designs and concepts to be launched as often as needed—even daily. With an advanced CMS, teams can test out new and creative approaches that can win over more customers, launching the company’s dynamic digital transformation.

But there’s a challenge associated with tailoring your site to serve each person and their unique requirements: archiving all of the personalized journeys in a clear and defensible manner. If your team must be able to show what your customers saw when navigating your site and what statements or promises you’ve made, advanced CMSs create [challenges for the compliance teams](#) who are responsible for [meeting regulatory obligations](#) to capture, archive, and preserve a live recap of their digital marketing and communications.





# REGULATORY COMPLIANCE ARCHIVING—HALFWAY IS NO WAY

Regulatory compliance is like pregnancy: you can't do it halfway. Either you're compliant or you're not. The bottom line is whether your archiving solution can collect and safely store everything you need? Can it keep you in compliance? Don't settle for a solution that only gets you halfway there—because partial compliance is just another name for noncompliance.

As cool as the websites created by advanced CMSs are, companies that use those systems find that they aren't all sunshine and roses. There's a fly in the ointment: companies in highly regulated industries such as finance and insurance still have to maintain archives to meet their regulatory compliance obligations. Worse, most of the traditional tools for online archiving—such as PDF capture and screengrab images—are ineffective at capturing the user experience within an advanced CMS. In some cases they fail outright, foiled by dynamic web applications that they simply can't manipulate or by personalized web pages that they cannot access.

This creates a scenario where your marketing team has its foot mashed firmly on the gas, building a dynamic, engaging modern website that satisfies heightened customer expectations, while your compliance team has its foot pressed equally firmly on the brake, with an outdated and inadequate archiving solution that restrains your creativity in service of your regulatory obligations.

It doesn't have to be that way. Here's where traditional archiving solutions go wrong.



## ARCHIVING REQUIREMENTS FOR THE FINANCIAL INDUSTRY

Regulatory compliance is only as good as the proof of that compliance, which is why recordkeeping is so important. You can communicate fairly and openly, provide all the right disclaimers, and make well-researched, specific recommendations that are suitable to your customers, but if your records don't establish those actions, you may find yourself figuratively up a creek.

Those recordkeeping requirements extend to statements made online, including on a company's website and social media. [FINRA Regulatory Notice 10-06](#) specifically requires that any firm communicating with the public online retain records of those communications that relate to its "business as such." [FINRA Regulatory Notice 11-39](#) clarifies that "[w]hether a particular communication is related to the business of the firm depends upon the facts and circumstances [of the communication]. This analysis does not depend upon the type of device or technology used to transmit the communication."

[FINRA Regulatory Notice 17-18](#) extended that rule to other forms of communication and noted that before a firm can use a method to communicate with customers about its business, it "must first ensure that it can retain records of those communications as required."

Links are potentially problematic too. Regulatory Notice 11-39 states that "Firms may not establish a link to any third-party site that the firm knows or has reason to know contains false or misleading content. A firm should not include a link on its website if there are any red flags that indicate the linked site contains false or misleading content."

Additionally, [Securities Exchange Commission \(SEC\) Rule 17a-4\(f\)](#), which allows and governs the storage of books and records in electronic format, notes that those records must be "preserve[d] exclusively in a non-rewriteable, non-erasable format (WORM)."



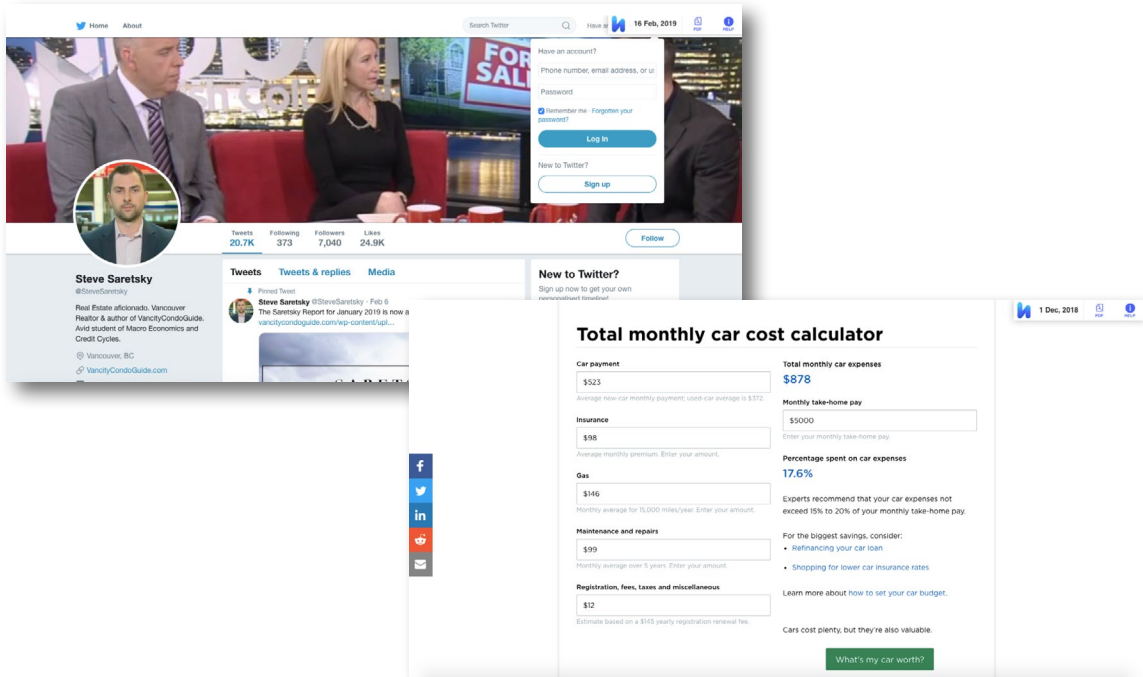
# WHERE LEGACY ARCHIVING SOLUTIONS FALL SHORT WITH ADVANCED CMSs

Traditional methods for creating website archives primarily rely on PDF or image file screengrabs. While they may be created by manually navigating a site—introducing the considerable possibility of human error—they're more often generated from an automated crawl that may (or, more likely, may not) be supervised or verified. We hear from a lot of clients who started with another solution and experienced one failure or another that led them to seek out a better approach. Typically, it's the organization itself, not the vendor or supplier who was supposed to be producing their archives, that spotted the error.

There are three key issues that make archiving modern websites—and especially websites created using advanced CMSs—challenging. First, archiving solutions that create static images fail to capture the dynamic, interactive content that these websites rely on. Second, most archiving methods miss out on the many individual permutations these websites offer, collecting—at best—one version of the site and missing out on myriad others. Finally, legacy archiving solutions generally require significant maintenance from the customer, which negates many of the advantages of using an advanced CMS.



Company C had been using another archiving solution prior to redesigning its entire website with an advanced CMS. After the new site launched, the customer checked its archives and discovered that absolutely nothing from the new site had been collected. Fortunately, the customer was proactive about trying to verify its archives and identified the problem quickly.



## PROBLEM 1: DYNAMIC CONTENT

Screen capture methods, whether they generate PDFs or image files, create static representations of websites that are anything but. These still images miss out on interactive elements like mouse-over or hover-over content and drop-down menus as well as custom information from fillable forms, calculators, and other complex tools. Each individual page may have a tremendous amount of data that's lost in a static image, which often includes disclaimers that only appear when a user interacts with certain page elements. Every one of the dynamic elements that make modern websites so user friendly is entirely non-functional in a PDF or a still image.

The only way to show how a user experienced a site at the time of capture is to show how it actually worked at that time—and static screengrabs don't do justice to your website experience.

For example, say you want to prove your website's compliance with the Americans With Disabilities Act (ADA). Advanced CMSs make ADA compliance easier than ever to achieve, ensuring that functions such as screen readers and alt text work across the entire site. But if all you have for your site archive is a static PDF, how can you prove that it was compatible with a screen reader?

This one problem would be bad enough to render traditional archiving methods obsolete, but it's only the beginning.



Financial Institution F uses dynamic graphs on its website to show the individual funds it offers, along with those funds' performance on specific dates. Each of the 200 available funds displayed on the page includes more detailed information and a specific disclaimer, but that content is only visible when the user hovers over the name of the individual fund. Capturing a PDF of each individual page for all 200 funds would be incredibly labor-intensive. These enhanced functions look terrific in development and enhance the end user experience, but they don't play well with legacy archiving solutions.

## PROBLEM 2: PERSONALIZED CUSTOMER JOURNEYS

When your website is highly customizable and changes based on who is browsing it, no one is having a single “typical” experience there. That means you can’t just create “an” archive, like you might do for a simpler site, and call it good (or compliant). Instead, the journeys your individual customers have depend on their choices—meaning that you need your archives to convey any, and every, potential customer journey. At a minimum, you should capture the major choice points and themes in those various journeys so that you can establish your compliance with any regulatory requirements.

Traditional archiving solutions miss the mark again here. They often cannot even access multiple versions of a webpage, such as A/B versions or location-specific iterations. While advanced CMSs allow variations in pages based on an individual user’s cookies, reflecting the number of times they’ve visited the site and the amount of time they’ve spent there—among other user characteristics—most archiving services only capture the static image of the page that loaded for them.



### Got omnichannel delivery?

If you’re using an advanced CMS to manage and scale omnichannel delivery across different devices and channels, you need an archiving solution that can access each of those channels and platforms—so that you can demonstrate your users’ experiences from any device, anywhere.



### PROBLEM 3: LABORIOUS SITE MAP MAINTENANCE

Most traditional archiving solutions require that the site administrator provide an up-to-date site map so that the vendor or tool knows what pages to capture. Providing a site map requires keeping track of and listing out every single page that should be captured. One-time provision of a site map wouldn't be prohibitive, but to create a complete archive, the site map must be updated every time a page is added. With advanced CMSs, pages or page versions may be added on a daily basis. Since the CMS manages itself, keeping pages organized and consistent and empowering your content team to publish new content freely with a minimum of busywork, providing those continual updates and maintenance for the sake of the archive crawler negates one of the major benefits of using an advanced CMS.

Often, we hear from compliance teams who have suddenly discovered that their archives aren't capturing their full dynamic site experience. Sometimes, they find that their archives haven't captured anything at all beyond a login screen. Or they discover that they're only archiving a single customer experience, when there are dozens of possible experiences that they're missing out on. We've also heard from customers whose previous archiving tool required multiple logins or preliminary steps—the very type of effort that an advanced CMS promises to relieve.

We think archiving requirements shouldn't get in the way of creativity. Your business shouldn't have to compromise on marketing innovation to remain compliant with whatever regulations you're subject to.

**Here's how to get the best of both worlds.**



# BEST PRACTICES FOR COMPLIANCE ARCHIVING WITH ADVANCED CMSs

What you need is an archiving solution that can automatically capture and securely store a fully functional version of any web content your marketing team can cook up. Look for web archives that are generated as completely navigable, interactive replica websites, preserving all the dynamic elements of the live site in secure, SEC-compliant WORM storage. These archives should be based on customized, in-depth web crawls that capture every aspect of each personalized and customized website, backed by an intensive and ongoing human QA process.

## VIEW FUNCTIONAL DYNAMIC CONTENT

By using the WARC (Web ARChive) file format, you can create operational website archives that demonstrate exactly what customers experienced on the live site at the time of capture. View all dynamic content, from videos and audio to rotating carousels. Engage with the interactive content of the original site, such as hover-over capabilities and fillable forms. With a powerful web crawler, you can scope any customer journey available on your live site and retrace your customers' potential steps within your archive to demonstrate your compliance to regulatory officials.

This produces a major competitive advantage over static captures, especially when it comes to proving ADA compliance. Unlike a PDF or image file capture that cannot demonstrate compliance with screen readers, a WARC archive can be navigated using many screen readers. Additionally, to verify that all of your pages have any necessary ADA compliance features, you need an archiving solution that can search its own content and create alerts that will immediately notify your site administrator.

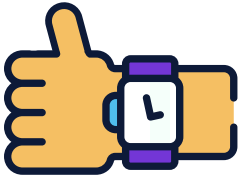


If you're like most marketing teams, you probably aren't confining your online efforts to your organization's own web page. That's why you need an archiving solution that also collects and archives dynamic content from YouTube channels, Twitter feeds, and any other social media sites that you use for your business.

## CAPTURE EVERY VARIANT OF A PERSONALIZED WEBSITE

With a personalized site, capturing the entire customer journey—in each of its potential iterations—is an important component of establishing regulatory compliance for every visitor. When there are several steps within a process, such as the process to obtain product offering or pricing information, companies must establish compliance with each branch along the path that leads to any given final data point. For example, you might need to demonstrate the granular options that could make a difference to the pricing and available options that a prospective customer saw based on their zip code, gender, or age. Otherwise, there's the risk that your archives will miss some permutation or skip a critical disclaimer.

Suppose an insurer that offers Medicare plans wants to demonstrate what plans it offered to a particular visitor and how those plans were priced. Both availability and rates depend on a range of variables, including the insured's age, sex, zip code, and any supplemental or additional coverage. Your web archiving solution should be able to customize its crawls in accordance with those, or other, parameters. This customization allows the crawler to capture every desired version of a site by accessing it from distinct IP addresses and viewing location-specific details. By preserving the site's full interactive function with the WARC format, other variables can also be manipulated within the archive to demonstrate every possible offering and its corresponding price.



## SAVE TIME AND LET YOUR TECHNOLOGY DO THE WORK

You shouldn't have to waste time constantly updating your site map. You need an archiving solution that can figure out when your marketing team has added a new page (or ten) or removed a page (or ten). Your archiving solution should have a sophisticated web crawler that can find all new pages and capture them without any extra effort on your part. And wasn't that the point of adopting an advanced CMS in the first place?

Look for an archiving solution that automatically discovers links within a parent page and crawls all of the linked webpages, creating dynamic captures of each. That way, you need only provide the parent site that the crawler should start from and those pages where there are different versions, such as A/B test options. A sophisticated crawler can access numerous different scenarios, log in with different IP addresses, and select various combinations of characteristics to reveal every possible version of a page for archiving. Once you've configured your initial capture, you're free to "set it and forget it," letting your CMS do the work of managing your content without constantly monitoring your site map.

Of course, if you're going to set it and forget it, you need to know that someone's watching, which brings us to our next point.





## HUMAN QA

Everyone wants to be able to “set it and forget it” with their archives, but let’s be real: it’s not always the most responsible decision. Someone needs to confirm that you’re actually capturing everything that you’re supposed to, right?

Absolutely. The question is just whether that someone needs to be you. That’s why your archiving solution should be accompanied by an extra layer of professional service: a human QA team who will verify that your collections are complete and functional. This requires that your archiving vendor monitors every collection to confirm that the basic metrics for your collections line up. This should include checks for whether the dynamic elements within a site—dropdowns, tabs, float-over content, and more—were all captured, as well as a confirmation that the total number of pages captured comports with the expected inventory. Ideally, your archivist will also conduct spot checks for functionality and periodic side-by-side comparisons between the live site and the archived version. Of course, if there’s ever a problem, you should hear about it from your vendor, instead of having to look for it yourself.

If you’re tracking your website visitors with your advanced CMS, don’t worry: you can ask your archiving vendor to exclude the archiving crawl visits from your legitimate traffic so you’re not muddling your results.

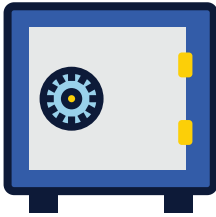


## CUSTOMIZED ALERTS AND NOTIFICATIONS

Need to verify that every page includes a general disclaimer? How about an SEC disclaimer? Want to know when a page is missing accessibility features for ADA compliance? Or when a page is missing an [opt-out button such as the one required by the California Consumer Privacy Act \(CCPA\)](#)?

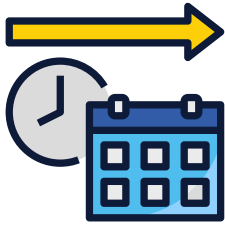
Look for an archiving solution that can create a customized alert for whatever you need to know about your captured website. You may need a capture notification to see what areas of your website are broken or where you have orphaned pages. We've seen customers ask for daily alerts whenever their archives contain particular phrases (such as "guaranteed results") or lack particular language (such as a required disclaimer).

There's another problem with traditional archives that we didn't even mention: they don't always allow powerful searches. One company explained to us that it wanted to search its site for specific disclaimer language, but its previous archive provider never returned consistent results. We recommended that the company add a notification feature. Now, if any pages appear without necessary disclaimer language, the company's site administrator gets an email alert pointing out those deficient pages for prompt resolution.



### **Secure, SEC-Compliant Storage**

Your archiving solution may not be able to guarantee the security of your website, but it definitely needs to protect your archives. Your captures should be stored in encrypted write-once-read-many (WORM) storage—as required by the SEC—through a secure, trusted web provider. As noted above, your archives should be created using the ISO standard Web ARChive (WARC) format, which works on any platform. You'll also want to verify that your archivist conducts ongoing audits of its security practices, and you may want to work with a vendor that has a SOC 2® Type 2 security certification.



## FUTURE-PROOF YOUR DYNAMIC DIGITAL TRANSFORMATION

If you have a complex, dynamic website—whether it’s created using an advanced CMS like Adobe Experience Manager, Sitecore Experience Manager, Drupal, or some other tool or combination of tools—you need to ensure that you also have dynamic web archiving capabilities.

With fully functional archives, generated from sophisticated web crawls, created in the WARC file format, and backed by human QA, you can future-proof your dynamic digital transformation. Whatever you can come up with for your website—today or in the future—you need to know that your archiving solution can capture it. When your archives are fully functional and interactive, including every customized and personalized variant, you can create dynamic, user-friendly websites without feeling constrained by regulatory compliance requirements. Incomplete static captures, missing versions of sites, constant verification and stress, and the need to update your site maps are all in the rearview.

Advanced content management systems give your marketing team freedom. Don’t let your archiving solution constrain that freedom.



Hanzo provides modern ediscovery and compliance software for enterprise organizations. Our solutions empower legal and compliance teams to efficiently manage the preservation, targeted collection, and review of dynamic content from enterprise collaboration applications, social media, and complex websites.

- **HANZO DYNAMIC CAPTURE** is the gold standard software technology providing high-fidelity, full context dynamic collection, archive, and preservation of dynamic web and social media content for both discovery and compliance needs.
- **HANZO HOLD** is the first purpose-built ediscovery solution to give governance and legal teams precise control over retention, preservation, and contextual early case assessment over hard-to-manage collaboration data.

Hanzo is SOC 2® Type 2 certified, demonstrating our commitment to data security, and serves large corporations across the globe—giving them control, visibility, and context for their data to reduce cost and mitigate risk. Learn more at [hanzo.co](https://hanzo.co)