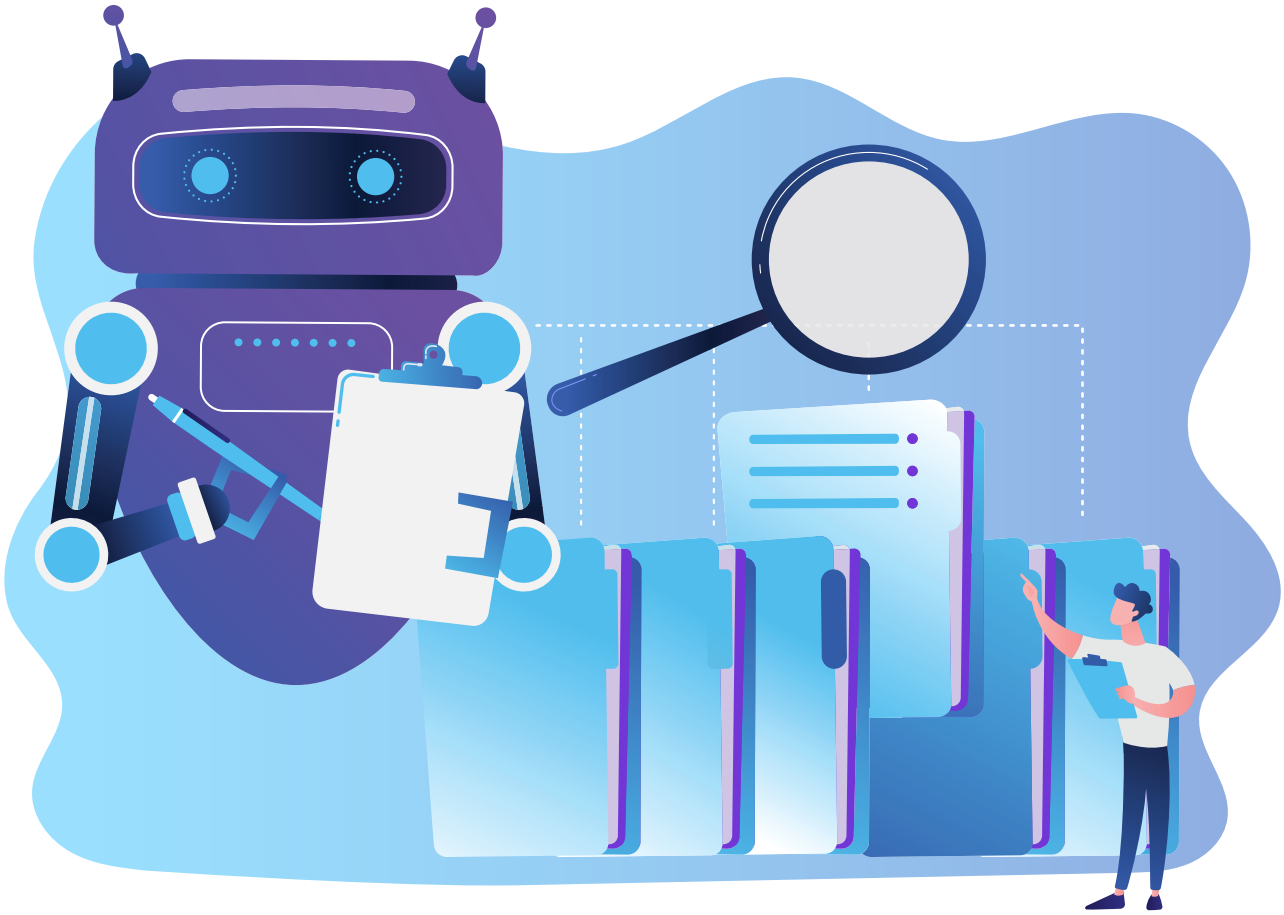
COMPLETE COLLECTION

Hanzo Web Archiving Essentials for Compliance

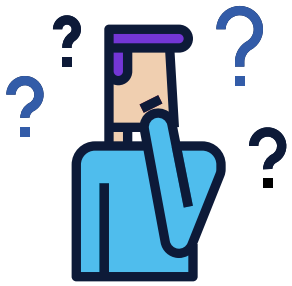
hanzo

HANZO WEB ARCHIVING ESSENTIALS FOR COMPLIANCE

4 Questions to Consider When Evaluating Archiving Providers



hanzo



4 QUESTIONS TO CONSIDER WHEN EVALUATING WEB ARCHIVING PROVIDERS

How confident are you in your web archive's fidelity? If you visit the websites of Hanzo's competition, you'll find quotes from their senior leadership touting them as "the financial services industry's premier electronic communications archiving provider." This blanket statement is, of course, carefully crafted to give you confidence. From our perspective, though, it's actually doing one of three things:

1. Misleading you about their technology's superiority to win your business.
2. Demonstrating how naive they are to the fact that their archiving approach is outdated.
3. Confusing brand recognition with customer value.

Which one is worse? At Hanzo, we want to ensure that you have the knowledge you need to make the right decision. Instead of making a sweeping blanket statement to do that, we'll show you, and tell you, what you should look out for. Whether or not you choose to work with us is ultimately your decision, but our goal is to give you the information you need to evaluate the archiving technology vendor landscape so you can make the right choice for your organization.

But first, a little background. For the past decade, Hanzo has helped legal and compliance teams in Fortune 1000 organizations solve their regulatory and risk-management web archiving challenges with a unique "native format" technology that has yet to be copied or replicated by our competitors. When the world's biggest brands, like Coca-Cola, speak with Hanzo's team of web and social media archiving experts about a problem, they leave that conversation confident that we can, and will, help them solve it. We built our technology because the solutions that already existed weren't good enough to properly capture the increasingly interactive, personalized, and layered nature of websites and social media channels.

> **"Hanzo was born out of a British National Library project to preserve the web forever—a daunting mission. Our co-founders were tasked with recording a contextual online experience for British citizens to look back on and analyze. There was only one problem when they were thinking about this challenge: every off-the-shelf tool and utility at that time either rendered a static flat PDF or extracted text into some kind of garbled and unreadable form, removing almost all of the dynamic nature of the web." — Keith Laska, Hanzo Chief Executive Officer**

WHEN CONSIDERING ANY ARCHIVING TECHNOLOGY PROVIDER, WHETHER YOU WANT TO CAPTURE AND PRESERVE YOUR WEBSITE, SOCIAL CHANNELS, OR INTERNAL COLLABORATION TOOLS AND TECHNOLOGY, THERE ARE 4 QUESTIONS TO KEEP IN MIND:

1. What technological approach is the vendor using to actually capture, archive, and preserve digital experiences and communications?

2. Does that approach meet the expectations set out by the regulatory bodies governing your organization and mitigate risks in alignment with your risk appetite?

3. Is the technology you buy today still going to be an effective solution in the future, or will it be outdated before your contract expires?

4. Do you trust this technology provider to continue helping your organization beyond the point of sale and implementation, delivering a level of customer service and account management that will reduce your stress and workload?

FUNDAMENTALLY FLAWED AND RISKY ARCHIVING APPROACHES: APIS AND PDFS

When archiving and preserving a website or a social media page, most vendors generally take one of two approaches:

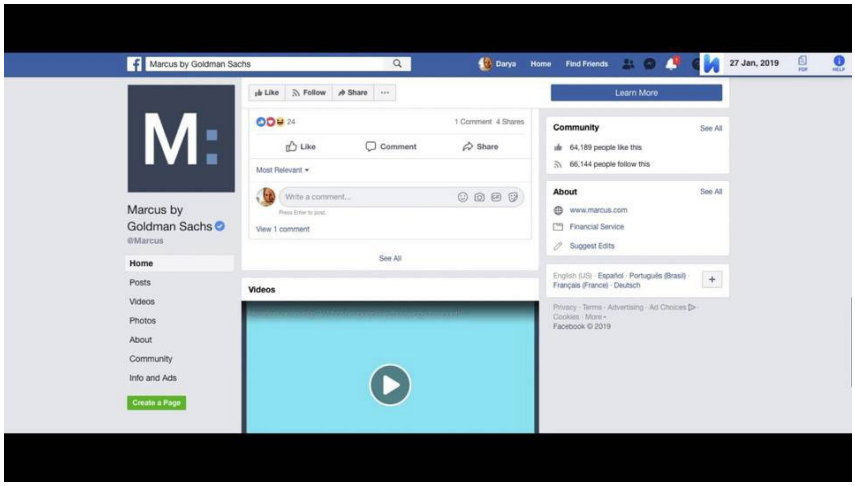
1. They use the platform's Application Programming Interface (API) to hook their archiving technology into the back-end of each social media network.
2. They take a dressed-up version of a static "screenshot" image and save it as a time-stamped PDF file.

Both of these approaches have serious flaws that have grown more obvious as the web has grown more complex. If you're responsible for allocating part of your annual budget to procure technology that will solve your organization's web archiving challenges, you have to ask yourself one question: do I have the risk appetite to buy technology that won't work all the time and could expose my company to regulatory failure and further scrutiny?

The biggest and most dangerous problem with relying on APIs to archive web and social content is simply that they leave you at the mercy of the API provider. An API might change for countless reasons—and when it does, technology that relies on that API will "break" until it can be re-engineered. A prime example of this took place in 2018, the year of data privacy awareness, GDPR regulation, and the Cambridge Analytica scandal.

> “With all of the recent attention on how companies gather, mine, and use personal data, many platforms have disabled or limited their API access to control external access to private user data. Facebook announced in April that it would limit the user data that developers could access; it then cut API access further in July. Instagram’s API access was similarly limited, as was Twitter’s.” — Evan Gumz, Hanzo

The truth is that Facebook, Instagram, LinkedIn, Twitter, and your awesome new website cannot be translated into a piece of paper. When vendors that use this approach originally introduced their archiving technology, it was primarily used for email and a more basic, text-based web browsing experience. Since then, the nature of the web has changed, growing more interactive, video-based, and dynamic—but the technology they want to sell you has, at its core, stayed the same. This creates unacceptable levels of risks for your organization.



First and foremost, if you try to translate your website and social media presence to a static piece of paper, there is a 99.9 percent chance that all of the content on those pages will not be captured, archived, and preserved. Right off the bat, from day one, the fidelity and accuracy of your archives are compromised.

Next, the context in which those messages, communications, and content were shared with your customers has been stripped away in the conversion to a static image—which FINRA, the SEC, and other regulators really don't like.

Lastly, precious data and metadata found within the website and social media pages you're trying to capture are likely to suffer one of two fates. They may be lost entirely in the translation from digital to PDF, leaving you with an incomplete picture of the facts. Alternatively, they may be easily tampered with, rendering them inadmissible as evidence in court.

All told, these two fundamentally flawed archiving approaches put your organization at risk of both not capturing and preserving everything you intend to and losing essential context and data within the files that you actually do capture.

Hanzo's Executive Chairman, Kevin Gibson, summarized these problems quite eloquently in a recent article he wrote for Artificial Lawyer:

> "Web data differs from every previous form of communication in that it is dynamic. Unlike email or Word documents, the internet has never been a paper-based medium; online data cannot be reduced to paper without losing critical information. The very nature of the internet is its interconnectedness and ceaseless changeability. The linkages between pages have inherent meaning, providing context, detail, and richness."

Hanzo's technology was built with the dynamic nature of today's internet experience in mind, as well as the regulatory and legal requirements around it.

REGULATORY REQUIREMENTS AND LEGAL DEFENSIBILITY: SEC, FINRA, AND ISO

We've covered some of the technical reasons why you shouldn't choose other web archiving vendors, but there are legal and regulatory reasons to consider too. In a perfect world, the data and content you archive and preserve will never have to be demonstrated to a regulator or used as evidence in an investigation or trial. But, unfortunately, we don't live in a perfect world, and compliance teams can't afford to operate in an environment where they're only circumstantially mitigating risks and complying with regulations.

Let's start with ISO 28500. Most compliance and risk professionals are familiar with the International Organization for Standardization, but perhaps not with this specific standard. Originally published in 2009 and updated in 2017, it establishes a standardized file format for the collection of navigable websites without any loss of information which might equate to spoliation of data and evidence.

There's a little-known fact about the WARC file and ISO 28500—Hanzo's founders were among those involved in creating and establishing this file type and demonstrating why it's the gold standard for storing, managing, and preserving billions of saved web pages in a universally recognized file format.

If you don't use WARC files, you could find yourself in the situation demonstrated by *Leidig v. BuzzFeed,* where the court deemed one party's archived content to be useless as evidence. Even better, the beauty of archiving web and social content with WARC files is that it brings you in compliance with FINRA and SEC rules and regulations, as we've detailed in our article on FINRA Regulatory Requirements for Archiving, Recordkeeping, and Supervision in 2019.

For example, under FINRA Regulatory Notice 11-39, a firm

> **"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."**

What this Regulatory Notice means is essentially that for any link your website makes to another website, you need to archive the linked page too, or you're at risk. A PDF may not capture the address for the link, much less the actual content of the referenced page. Additionally, the link in question could change, or the page could be deleted, removing essential evidence and data from your archive.

At Hanzo, we call these links “hops.” Anything we archive can include a “hop” to a linked page, and another hop after that. That allows you to:

a. follow third party and social media links;

b. archive linked pages in their native format at the moment they are referenced, in case they are deleted or altered in the future;

c. capture those pages in the context in which they are referenced, so the full story and experience of navigating to that piece of information, including how it was presented, is fully preserved.

A similar rule, and concern, applies to social media. FINRA Regulatory Notice 17-18 states that “By sharing or linking to specific content, the firm has adopted the content and would be responsible for ensuring that, when read in context with the statements in the originating post, the content complies with the same standards as communications created by, or on behalf of, the firm.” That means you need to capture “hops” and linked content on social media too. Now let’s turn our attention to the SEC, specifically SEC Rule 17a-4, which lays out the details and criteria for how electronic communications need to be preserved. It requires that:

1. Archived electronic communications are stored in a format that is non-rewritable and non-erasable, also known as WORM (write once, read many).

2. There is a way to verify the quality and accuracy of the archived content.

3. The archives are serialized, retained on electronic storage media, and downloadable to any other accepted medium.

Hanzo's web archiving technology, and the WARC files at the heart of it, check all of these boxes—because that's what they were designed to do. By contrast, the PDFs and other image file formats that other vendors use may not meet any regulatory requirements or may meet some but not all of those requirements. When exploring other vendors and weighing your options, make sure you ask whether they use WARC files. If they don't, ask them why. If they can't give you a good answer, you may want to reconsider whether they're the right fit for your organization.

TRUST, LONGEVITY, AND QUALITY ASSURANCE

We recently conducted an archive risk analysis for an organization working with another vendor. Upon completing that analysis, we discovered that the majority of their website was not being captured and archived, despite their expectation that it would be. The core of this problem wasn't any malicious attempt on the vendor's part to mislead the organization. Rather, the vendor took a faulty approach, relying on a sitemap to determine what to capture, and then magnified its error through a lack of quality assurance (QA) and testing. When it comes to risk management and compliance, it's always easier to prevent and prepare for what you know; it's the blind spots, and the false sense of confidence they enable, that can cause the biggest problems.

Here's the thing: sitemaps change, and frankly, they aren't always 100 percent accurate. Whenever a new page is added to your website, it complicates the existing sitemap model, which can result in missing content and data on both the front end and the back end. In other words, working with a vendor that relies exclusively on sitemap information without implementing a certain standard of quality control can result in days, weeks, or even years of missing data and unarchived content. It's also essential, over the long run, to be able to manage and search your entire archive with specific criteria, keywords, and Boolean operators that help you find relevant information.

Hanzo's crawling technology automatically discovers new web pages created by your team and adapts to capture those pages. It can also deliver alerts when certain keywords do or don't appear, and it's capable of searching for conditional information across millions of archived pages. During our onboarding process, we aren't just analyzing the accuracy of your crawl and determining whether it's capturing the pages you need. We're also measuring the fidelity and purity of the archived content against its live counterpart. This level of qualitative QA goes well beyond the numbers, ensuring that the content you capture is actually usable, and valuable, in the future.

Last, but certainly not least, we have the ability to customize our technology and write new bespoke code to meet a unique need or use case. For instance, we might create code to archive a very specific, personalized path on your website that someone would encounter if they provided certain information and criteria.

If you have a problem that isn't solved by our existing web archiving technology set, contact Hanzo today and let us know your specific need.

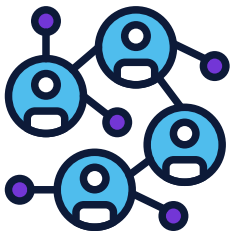
Archiving Instagram, LinkedIn, Facebook, Twitter, and Social Media for Regulatory Compliance

ARCHIVING INSTAGRAM, LINKEDIN, FACEBOOK, TWITTER, AND SOCIAL MEDIA FOR REGULATORY COMPLIANCE

When you need information, where do you go? For the majority of people today, the answer involves a smartphone and social media. From Facebook and LinkedIn to Instagram, Twitter, and even YouTube, social media is everywhere—and full of answers, data, opportunities, and potential risks.

For organizations, and the risk, legal, and compliance professionals within them, that means consumers are using your business's social media profiles and website to make critical decisions. If you're not keeping full records of what you're communicating online and on social media, you could be falling short of the regulatory compliance requirements imposed by the Financial Industry Regulatory Authority (FINRA) and the Securities and Exchange Commission (SEC), exposing your organization to easily avoidable risks and fines.

If your organization has an active social media presence, there's a strong chance you need to be archiving all of those communications, and possibly the communications your employees make on behalf of your organization too. For financial services businesses that provide information on social media, each post, comment, and share can be a business communication that is subject to the same regulatory compliance demands under FINRA—namely, recordkeeping and supervision—as any other form of communication. But archiving social media is harder than you might think, especially if you need archives that will stand up to a regulatory inquiry or hold up as evidence in an investigation, and it's possible that if you're currently using a PDF-based format to archive this digital content, you are at risk of non-compliance.



ARCHIVING SOCIAL MEDIA IS MORE COMPLEX THAN YOU MAY REALIZE

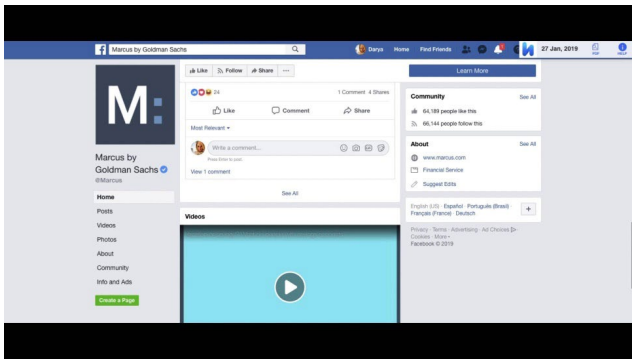
Think about what a typical social media feed looks like. It's ever-changing, complex, and interactive. It is decidedly not "what you see is what you get" and it's been designed that way on purpose.

To keep social media users interested, engaged, and actively clicking, social media feeds are set up to be interactive and expandable: there's always more information just a click away.

To capture that experience, the context surrounding it, and the data within it in a static PDF document is somewhat impossible. You might be able to see that a post has 18 likes and 4 comments, but without clicking on those links, you can't see who liked the post, who commented, and what those comments said, let alone who liked or responded to the comments. But why does that matter? Why do you need that level of granular, specific information, and why this urgency to preserve it?

From the perspective of a regulator, your compliance program is only as strong as your ability to prove compliance. When it comes to regulatory compliance, merely liking a comment, which may seem completely innocent and harmless, can equate to the adoption of its contents and endorsement of its message, potentially triggering additional recordkeeping and supervisory requirements.

Since Hanzo was founded in 2009, we've archived billions of pages in an interactive, dynamic format designed to solve many of the regulatory and legal challenges that PDFs present. With the rise of social media in the enterprise, we've become experts at capturing and preserving Facebook, Twitter, LinkedIn, Instagram, and other common platforms.
Here's a video demonstrating some of our social media-specific capabilities, taken directly from the content we've archived over the past year. It is a prime example of why PDFs aren't an adequate way to archive social media (or, really, any complex and dynamic online content).



HOW TO ARCHIVE SOCIAL MEDIA FOR REGULATORY COMPLIANCE

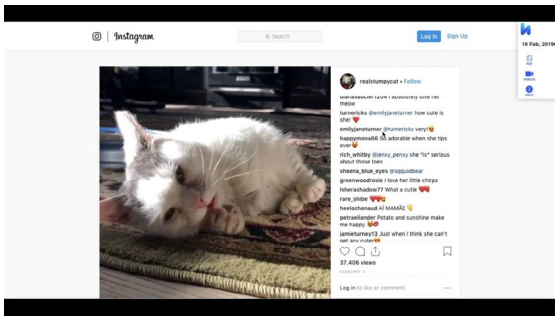
At its core, the web, including social media, is built around links, videos, and interactive, interconnected experiences.

The interconnectedness between sites is a key distinguishing feature of online communication, and while that's great for consumer engagement, it's replete with traps for the unwary when it comes to archiving and preserving digital experiences.

Whether you want to establish an archive of your social media presence and behavior to meet regulatory requirements, mitigate risks, or proactively capture and preserve essential data, here are a few key questions and facts to be aware of and keep in mind while auditing your existing archives or exploring potential technology providers.

1. **Is my web archiving solution designed to capture the full context around the communications my organization is sharing online?**

We've said it in the past and we'll say it again in the future. Context is essential to understand any piece of data or information, and context matters in litigation and compliance too. You need to capture the full dynamic context of all communications for your archives to be complete, and FINRA even has specific regulatory notices covering the importance of context. In the Instagram example below, taken from our existing archives, you'll notice a few specific things in relation to capturing the context around data:



On Instagram, similar to modern web design and other social platforms, images can appear in a "carousel." That means one post can contain layers of pictures and video, all related to each other. You'll also notice that the comments and activity on a particular post, when it reaches a certain point of "engagement" (often beyond 10 likes or a few comments) become hidden, only visible when expanded upon. From the perspective of a cat video, much of this seems harmless, but in relation to your organization, the information and data living in those "hidden" comments and likes, or layered within the video of your post itself, is extremely valuable and important for your own due diligence and investigations, as well as regulatory compliance expectations.

This one URL, archived in its native format with all of the functionality and data preserved as you see in the video, would take hundreds (if not more) of PDF pages to capture the full post, and navigating those pages would strip away the context of what the person using Instagram actually experienced. Consider this situation:

You work for a financial services company, and on your brand's LinkedIn account, the social media manager on your marketing team publishes a new, approved article with information on stocks and other financial investment opportunities that are "trending up" and trending down." While the article itself poses no compliance or regulatory risk and contains the necessary disclaimers to suggest the information itself is not advice, people who follow your organization on LinkedIn begin to share their own thoughts on the market in the comments section.

The 11th comment on that post expresses a specific interest in a specific stock, and someone within your company who reads that post and agrees with that comment, without thinking twice about it, likes their comment.

Your employee just unintentionally endorsed that trading suggestion and provided advice, but not just to that one person. LinkedIn's algorithm now shows their followers they liked that specific comment, unintentionally signaling to their community a certain sentiment about the financial outlook of that particular stock.

In your PDF web archive of LinkedIn, which contains a static screenshot of that post and only shows the first 3 comments, this behavior is virtually invisible, keeping your legal and compliance team in the dark to an actual compliance violation.

As you can see, dynamic capture is hard enough, but financial service providers need even more from their archives, increasing the technical challenge of social media archiving.

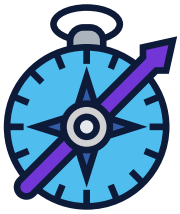
2. Are my web and social media archives quality tested by experts and preserved in a robustly future-proof format?

Your legal or compliance team might be retaining archived digital experiences and the data they contain for just three to five years, or you might find that you need them for 10 years or more. With the ever-accelerating rate of change in technology, the last thing you need is to turn to your archives a year from now, or 10 years from now, only to learn that you can't access them!

Under SEC Rule 17a-4, your archives must be maintained in a non-rewriteable format, commonly known as "write-once-read-many," or WORM. This requirement ensures that archives cannot be changed or altered after capture—and it's a common stumbling block for regulatory compliance, as evidenced by recurring violations and fines. Photo and document editing technology has gotten so advanced and easy to use that manipulating the text or images in a PDF-document is as simple as downloading a free iPhone app.

But even if you're comfortable with the risk of non-compliance around WORM, or aren't worried about your archives being manipulated and tampered with to delete or compromise precious evidence and data, there's a level of risk around the quality of the information actually being captured if it isn't being tested, and vetted, by your internal team or the vendor you work with to solve these problems.

For example, you have a PDF or screenshot-based archiving system in place to capture your homepage on a daily basis, but one month after establishing that system, you redesign the homepage or change the order information appears, in order to promote a new product or service offered by your company. Suddenly, the screenshots captured and saved as a PDF are missing essential information or don't look as good as they did previously, and nobody realized until it was too late, and for months, you've been archiving and preserving an incomplete picture of your digital presence.



3. Are my web archives navigable and searchable by specific criteria and keywords?

To adequately supervise your communications, you need archives that are navigable and searchable. Collecting all of this data in the proper format and with the right context is essential, but you also need to be able to actually use that information when you need it without spending hours manually trying to find the right data. When you only capture a PDF or screenshot, you lose out on data, metadata, and keywords that exist on, and beneath, the surface of every digital experience.

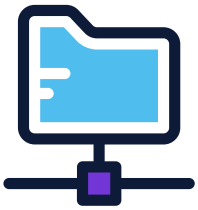
With Hanzo's dynamic website and social media capture, you can navigate and search an archived site the same way you would interact with a live site. You can scroll down a feed, clicking on links, expanding comments, playing videos, and even interacting with fillable forms and calculators. You can also look for specific keywords or pieces of data across thousands of archived pages to find exactly what you're looking for.

With Hanzo Dynamic Capture, you can show who liked a post or a comment, establishing that your business communications have been proper. You can also explore the full context of any statement so it can't be misinterpreted or taken out of context. Our archives are captured and maintained in a future-proof Web ARChive (WARC) file format that works on any platform and any system, today and at any time in the future. The WARC file format is backed by ISO standard 28500 and is the file format preferred by the Library of Congress. Hanzo's archives are also captured and maintained in a WORM format that complies with SEC Rule 17a-4.

Because they're fully navigable, our archives are also searchable, making them easy to review both for eDiscovery and for regulatory supervision, and because our work centers around the needs of litigation and compliance professionals, we've built our archives to be both admissible and subject to authentication. When you need to archive your online and social media communications, don't waste your time and budget on vendors and technology providers that don't understand what it takes to create defensible, workable website archives for regulatory compliance. Ask for a demo, see what they show you, and remember the three questions we've posed in this article.

The potential risks of non-compliance are easy to avoid, but it starts with making the right decision about who to work with. To learn more about how Hanzo can help, schedule a free consultation to discuss Instagram, your website, or other social media channels used by your organization today.

WARC and WORM Digital Storage



WARC AND WORM DIGITAL STORAGE

WARCs? WORMs? Is this a lost installment of the Lord of the Rings? Unfortunately, no. Rather, WARC and WORM are abbreviations that you should be familiar with if you're maintaining digital archives for regulatory compliance, especially in the financial services industry. WARC is a file format for web archives, while WORM describes a type of storage that's protected from overwriting. Let's break these terms down a bit more and review why they're important for archiving online communications and messages that live on your website and social media channels.

WARC, logically enough, stands for Web ARChive, a file format that fully captures the content of a website. Why is capturing the entire content of a website important? That's a two-part answer.

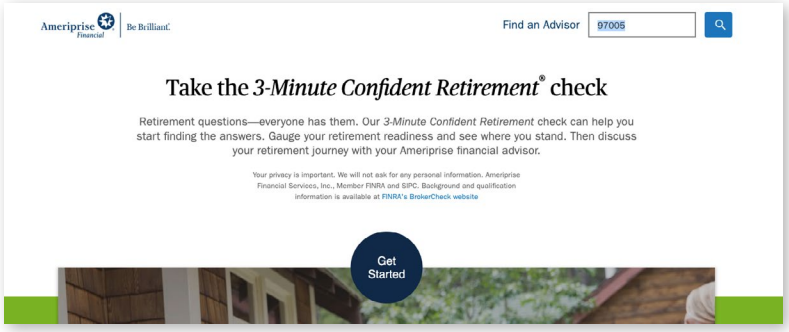
For starters, both the Financial Industry Regulatory Authority (FINRA) and the Securities and Exchange Commission (SEC) require that you create and maintain archives of all your business communications, regardless of where they occur.

The supervision and record-keeping provisions of FINRA and the SEC explicitly include online communication: FINRA Regulatory Notice 10-06 extended the standard business records requirements to communications that occur via social media, while FINRA Regulatory Notice 17-18 clarified that those requirements apply to messages sent via text or chat applications as well.

The bottom line is that all customer communications, whether they occur on Facebook and LinkedIn, over email, on your website, or another digital platform, need to be retained for compliance oversight.

Second, when archiving online content, you need to capture all of it, not just the "easy" parts.

Sure, you could take a screenshot of your website or save it as a PDF. That might look similar at first glance and might (or might not) display all of the words that are on the screen. But your website is almost certainly more complex and dynamic than a static image could ever capture. That's because typical website design includes interactive components such as introductory videos, photo and text carousels, and fillable calculators.



The Ameriprise website is a good example: its menu options are interactive, only appearing via mouse-over selection, and it includes a "confident retirement" tool (shown below) with choose-your-own-adventure buttons, where the accompanying text changes depending on your selection.

Will a still-frame PDF adequately convey all of the communication happening on that website? No way. It's even worse with social media; how can you prove that the company itself didn't like a client's positive comment on its post, thereby adopting that comment, if you can't investigate, within your archive, who responded and how?

For that matter, how can you supervise the activity of associated persons if you can't navigate your archives to explore the full context of a conversation? Oh, that's right. You can't. This is where the WARC file format comes into play.

WARC files are generated from a web crawl, in which software "crawls" through every link and component on a webpage and downloads that content along with its structure and description.

WARC files also download full supporting metadata, which allows confirmation that the archive was captured in its original format on a specific date. Each component on the page is captured in its own WARC file, which specifies not only what the content should include but also what it should look like and how it should respond to user interactions.

Those files can then be reassembled, creating a replica website that looks and operates exactly like the original site did. Hanzo will let you directly experience a WARC-file web archive and take it for a test-drive.

WARC archives can also be accessed from any platform and any operating system, and they're future proof. They'll work just as well in five years or 10 years or even 30 years, which is critical when you have to retain records for a decade or more.

How do we know WARC-based archives will stand the test of time?

Because the structure and function of WARC files are memorialized in ISO standard 28500:2017 and maintained by professional archivists. They're the archival format used by institutions that are in the business of maintaining records over the seriously long term, like the Library of Congress.

> While WARC files aren't technically required by FINRA or the SEC, they represent the best way to create compliant archives that enable full supervision of online communications and comprehensive record-keeping.

For a more in-depth technical look at how Hanzo uses WARC files, check out our WARC FAQ. But there's another component to fully compliant web archives, and this one isn't optional: WORM storage.

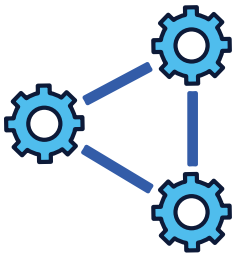
WORM—which stands for "write once, read many"—describes a type of file storage. Refreshingly, it means exactly what it says it means: WORM storage devices can be written on only once, but then their contents can be read many times. If you remember the early days of CDs, before they were generally rewriteable, you've experienced WORM storage. You had to figure out every song you wanted on your "1995 Summer Beach Music" mixtape before you started burning the CD—or else you'd have to start over with a new disc. Similarly, but with less Hootie and the Blowfish, the traditional tape archives used to back up business data operated in a read-only format that couldn't be overwritten.

As we mentioned, WORM storage isn't just a good idea: it's a requirement under the SEC's Rule 17a-4. In subsection (f)(2)(ii)(A), the rule notes that records that are stored on "electronic storage media" (rather than microfilm or microfiche) must be "preserve[d] ... exclusively in a non-rewriteable, non-erasable format." This isn't surprising, since the purpose of an archive is to ensure that a correct record is maintained. If an archive could be overwritten, edited, or "cleaned up" after the fact, like a PDF file, it wouldn't be much of an archive. Most of the digital storage we encounter today is rewriteable storage, though, such as USB thumb drives, internal or external hard drives, rewriteable or "R/W" CDs, and typical cloud storage repositories like Dropbox. The data on these storage media can be readily written and overwritten until the media itself becomes corrupted by overuse.

That means these commonly used storage devices aren't appropriate for archiving business communications. At Hanzo, we've designed our software and services specifically around the needs of regulatory compliance and eDiscovery professionals. All of our web archives are created using WARC files and stored exclusively in WORM storage, because that's what it takes to have robust, complete, and regulator-ready archives that stand the test of time. Our archives are, thanks to the WARC format, fully navigable, making regulatory supervision and eDiscovery review a snap, and they're readily admissible in litigation, which means they pass the sniff test for regulatory agencies as well.

Ready to explore the best way to create compliant archives? Contact a compliance expert at Hanzo.

The Importance of Capturing Context for Regulators



THE IMPORTANCE OF CAPTURING CONTEXT FOR REGULATORS

Context is essential to understand any piece of data or information. "Those are some gorgeous melons!" could be someone's positive feedback about produce at a farmer's market, or sexual harassment directed toward a farmer.

The news proves every day that out-of-context remarks can be tremendously dangerous, especially in our online environment today, where any piece of content can be captured, edited, and shared on social media to millions of people with ease. Statements taken out of context can bring the most innocuous celebrities to their knees, wreck politicians, and create PR nightmares for brands. The question always remains: were people's remarks actually taken out of context, or did they mean exactly what it sounded like they meant? Those lingering doubts can make it hard to overcome context errors in day-to-day life.

Context matters in litigation too. Take, for instance, a defamation case, in which ABC's 20/20 featured allegations that a pastor had misused church donations. The pastor could be seen clearly stating, on video, "I live in a 25-room mansion. I have my own $6 million yacht. I have my own private jet, and I have my own helicopter and I have seven luxury automobiles." Sounds awful if you're a dedicated church donor, right?

But as his lawyer said, "If you take something out of context, you can defame some[one] just as easily as if you make up the words yourself." The catch was in the context: the full video evidently showed that the pastor was speaking not about himself, but from the hypothetical perspective of a financially well-off, yet spiritually bankrupt, person. Those same context errors can be fatal in regulatory compliance. Unfortunately, common web-archival methods that rely on static media like PDF files don't fully capture the dynamic context of live internet or social media communications.

Regulatory compliance demands a clear view of customer communications, including their context. Whether it's in an internal corporate investigation into alleged misconduct or a full-blown regulatory action involving customer communications, it can be hard to know what someone said or meant without seeing the environment where the statement unfolded. It may be that a personnel action is warranted, or it may be that innocent actions have been misconstrued.

Proving what happened—as with the pastor in the video—can be challenging if you don't have a fully navigable record of online conversations. For example, under FINRA 10-06, the determination of whether a particular communication is actually a "recommendation" for purposes of Rule 2310 depends on the facts and circumstances—the context—of that communication.

> More broadly, FINRA Regulatory Notice 17-18, Social Media and Digital Communications, clarifies that customer communications occurring online or over social media are subject to the same requirements of fairness, completeness, and truthfulness as other communications.

Firms must retain full archives of their communications about anything that qualifies as "business as such," regardless of the medium by which those communications occur. Regulatory Notice 17-18 also states that a firm can adopt content, and therefore make itself responsible for ensuring that the content complies with regulatory demands, merely "by sharing or linking to [that] specific content."

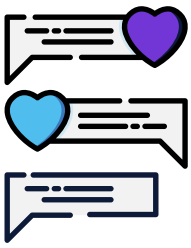
In other words, to maintain FINRA compliance—and to prove that compliance—you need navigable archives that extend beyond the target page, capturing related or linked pages as well. In a nutshell, you need context. And don't forget FINRA Regulatory Notice 18-15, Heightened Supervision, which requires that firms assess all of their associated persons, identify those with behavior that raises concerns, and develop heightened supervision plans where needed to protect investors. Keeping up with supervisory reviews demands the ability to review the context of statements to ensure that everyone is on the up-and-up and that those under supervision aren't unfairly judged.

The quick-and-dirty way to archive a website is to create a screenshot or a PDF. In, out, done—right? Not so fast. With PDFs, you're pretending that the internet can be reduced to paper—a static, one-dimensional representation of a media that is anything but static and one-dimensional.

> Static archives lose—or, at their very best, muddle—all of the rich context surrounding a statement. After all, you can't click a link on a piece of paper to investigate what content a company has adopted or endorsed.

Modern websites also rely on interactive JavaScript components to engage users and answer basic questions about the company's services. That means that, to demonstrate that those communications did not contain unfounded or unwarranted representations, you need the ability to fully play back any interactive content. You can't do that with a PDF!

Without navigating the full dynamic website and interacting with drop-down menus, chatbots, and animations or videos, you could easily find your words skewed or misinterpreted. For instance, capturing one moment in time—without the surrounding context of disclaimers and explanations—could make it appear that an investment company made a specific promise of financial performance. Exactly what that broker—or the company's web chat application—said to a customer could be the foundation for a costly and damaging compliance investigation. With a static archive, how can you prove that that one moment is being taken out of context?



The same goes for social media. Practically everything on Facebook, Instagram, LinkedIn, YouTube, and Twitter is interactive. Figuring out the context—who said what to whom, and who reacted to those comments—depends on having a navigable version of the full feed. Without that, you're left flipping through still images in an attempt to reconstruct events, which is slow, clunky, and not very compelling.

There's one more potential problem with static captures in this era of hacking, data breaches, and social manipulation: not everyone online is who they claim to be. Someone operating a fake profile using a business's name or an employee's name could wreak havoc by making inappropriate remarks. Without a full dynamic capture that allows the supervisor or the regulatory authority to investigate that profile and determine its authenticity, you'd have no way to prove that the comments weren't attributable to the business.

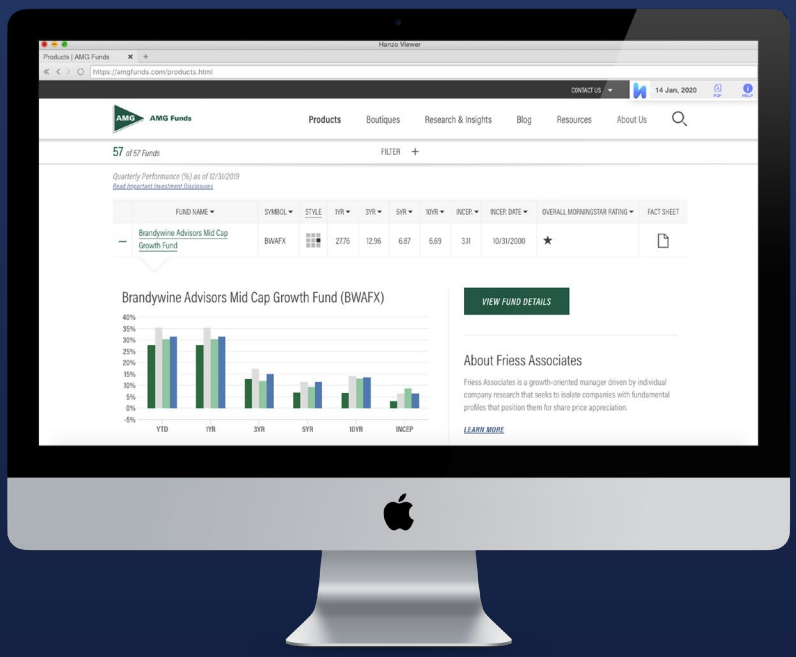
In short, to maintain regulatory compliance with your online and social media communications, you need to capture them in context, maintain that context in your archives, and have the ability to review and supervise communications in their original dynamic format.

That's why Hanzo recommends using native-format WARC (Web ARChive) files to generate fully navigable archives that look just like the original online content. With native-format archives, you can interact with any part of a communication, exploring the full context of all communications and proving exactly how a conversation unfolded. These types of archives also capture all dynamic content, including videos and interactive features, and links are fully clickable. That means you can investigate every interaction and every potential adoption or entanglement directly from your archive—and you can prove what did or didn't happen.

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 is SOC 2® Type 2 certified, demonstrating Hanzo's commitment to data security and privacy. The company serves large corporations across the globe — giving them control, visibility, and context over their data to reduce cost and mitigate risk. Learn more at hanzo.co.

DON'T LET COMPLIANCE STAND IN THE WAY OF INNOVATION

Now you can have all of the interactive, personalized, and dynamic web features available and still demonstrate regulatory compliance.



WE CAPTURE THE DYNAMIC WEB — IN ALL ITS INTERACTIVE GLORY — FOR YOU

Hanzo Dynamic Capture is a technology that can archive and preserve digital, social, and online forms of content and communications in an interactive, native and tamper-proof manner that can be easily searched and presented, while complying with industry-specific regulatory expectations.

REQUEST A DEMO

hanzo

Find out more: hanzo.co