



CASE STUDY

Case Study: Large-Scale Website Archiving for a Global Technology Company

Archiving 150,000 webpages to meet
the needs of an organization's legal and
marketing teams



Pagefreezer



Introduction

A global technology company headquartered in Silicon Valley posted an RFP that invited vendors to submit proposals for the website archiving of all its webpages—around 125,000 pages in total.

The RFP made it clear that:

- All 125,000 pages had to be **crawled and archived daily**
- The organization had to be **able to view and compare changes** that were made to webpages over time
- Since the company's previous archiving solution was reaching the end of its life, a new solution had to be selected and **implemented quickly**
- As a consumer-facing website with a significant eCommerce component, the site had a **complex structure that had to be captured and recreated exactly**, including navigational items like menus
- Any chosen vendor would have to illustrate thorough systems and processes for **keeping data secure**



Pagefreezer was quickly shortlisted during the RFP process, at which point more information was provided to applicants. The organization explained that it would largely be the legal and marketing teams using the archiving solution. They would need it for the following reasons:

1. eDiscovery and Litigation Readiness

Web content is a crucial part of modern business communications. It is also often part of essential legal documentation, including offers, promotions, purchase orders, and online transactions. So it's imperative that businesses keep such documentation to support lawsuits when they arise.

Successful global enterprises make for particularly attractive targets when it comes to accusations of false advertising and other website-related litigation (such as [non-compliance with the Americans with Disabilities Act \(ADA\)](#)—which was why the legal team needed accurate and defensible records of website content.

2. Monitoring and Managing Website Changes

In order to minimize the likelihood of website-related complaints and legal matters, the organization had clear processes for the review and approval of new web content. By allowing easy comparison between different chronological versions of the same webpage, a website archiving solution could provide some added peace of mind. If there were any questions about when and how a particular page had been updated, it would be easy to identify and review these changes.



Automated Website Archiving with Pagefreezer

After review, Pagefreezer was selected to archive the organization's web content. Key to this success during the RFP process was how **Pagefreezer simplified the collection, review, and export of website content.**

Pagefreezer captures and creates exact replayable snapshots of a website. This means legal and marketing teams can view videos and images, navigate via existing menus, and even access password-protected pages and user-generated content. Pagefreezer also captures the source code and style sheets of a website. So as an added benefit, a page or site can be rebuilt from Pagefreezer archives if it is lost for some reason.

Teams can also use **advanced search** to find what they are looking for across all pages on multiple websites—and then view that content as it appeared on the original page.

Additionally, the marketing team can use the compare function to quickly see how content has changed. So if, for example, the compliance department asks for a record of all recent changes to the company website, it's easy to see what content looked like on a given date—and how it has been altered over time.

Should there be a request to produce a piece of web content for possible litigation, Pagefreezer allows for the **easy export of evidence.** Data can be exported in various formats—complete with a **timestamp and SHA-256 digital signature** that prove authenticity. Pagefreezer data is also designed to be ingested by modern eDiscovery platforms, like Relativity.



Enterprise-Level Data Security

As mentioned earlier, the organization viewed robust data security as another non-negotiable. So, even if a solution met their needs, it would not be selected if it introduced any vulnerabilities.

Thankfully, Pagefreezer successfully passed this security review due to its impressive platform security and credentials.

Pagefreezer's product and organizational security are designed to ensure only authorized users gain access to your website archive. Pagefreezer offers **enterprise-level security features**, including single sign-on (SSO), two-factor authentication (2FA), IP whitelisting, concurrent login management, and password policy management.

Advanced user, group, and role management makes the appropriate provisioning of users simple and easy. The archive activities of all users are also logged to easily monitor actions. These **audit logs** give platform administrators detailed insight into all activities on the system, including what exactly was done, who did it, and when this activity took place.

Pagefreezer uses **network and performance monitoring tools** for ongoing monitoring of servers, systems, and applications to assess health and performance, availability, and capacity. Pagefreezer also has an **Information Security Incident Management Plan** in place to ensure that all events related to information security—or weaknesses associated with information systems— are quickly responded to.

Pagefreezer is SOC 2 Type 1 and Type 2 compliant. Our independent auditor's report attests that Pagefreezer has put in place controls for information security and confidentiality that are suitably designed (according to the trust services criteria), and that after in-depth testing and examination, these controls operated effectively throughout the review period. Data centers that we use in North America are also SOC compliant.

Pagefreezer's management system is ISO 27001:2013 certified, meaning that we consistently meet the security goals outlined in ISO 27001. This includes limiting data access only to those who are authorized, protecting data integrity by preventing unauthorized alteration, and offering customers reliable access to the data that they need. Data centers that we use in Canada and Europe are also ISO 27001 certified.

Would you like to learn about Pagefreezer's Web Archiving solutions?

Pagefreezer's website archiving solutions protect you from false claims, intellectual property theft, and freedom of information requests related to website content.

Visit our [Website Archiving](#) page,
or simply connect with one of our solution advisors:

[**Book a Demo**](#)

Email:

sales@pagefreezer.com

Phone:

+1.888.916.3999 (North America)

+44 (0)20 3314 7921 (U.K.)

+31 (0)76-5324275 (Europe)

pagefreezer.com