



Recordkeeping of Online Data in Federal Government

Managing websites, social media, and other dynamic data to meet FOIA and NARA requirements



CONTENTS

3

Introduction

4

The Implications of Directive M-19-2

6

Dealing with Increased FOIA Requests

8

The Need for Transparency

9

The Challenges of Online Data

10

Recordkeeping Example:
U.S. Department of Health & Human Services

Introduction

Record retention processes within federal agencies have received a lot of attention as of late. Most prominently, missing Secret Service text messages, White House call logs, and Presidential Daily Diary entries [related to the January 6, 2021 Capitol breach](#) made headlines and shined a light on the importance of reliable recordkeeping in the federal government. Similarly, in a letter to the House Committee on Oversight and Reform sent on February 18, 2022, [Archivist David Ferriero wrote](#) that: “The Trump Administration did not fully capture, and therefore NARA did not receive, all of the Presidential records created by Trump and White House staff that were posted on social media platforms.”

And speaking of the National Archives (NARA), Directive M-19-21 set out by NARA and the Office of Management and Budget (OMB) also reached its deadline in 2022, meaning that all federal agencies are now expected to conform to the stringent digital recordkeeping requirements outlined in the memorandum.

Lastly, there is the issue of ever-increasing FOIA requests. Federal agencies have experienced a steady increase in FOIA requests over the last decade, with 2021 seeing a total of 838,164 submitted requests. Needless to say, this kind of volume places agencies under tremendous pressure and demands new scalable solutions for the processing of records requests.

All of the above needs to be considered within the context of a recordkeeping environment that is evolving quickly. Gone are the days of paper records—and even discrete digital documents. Web modernization initiatives, migrations to cloud-based solutions, and the adoption of social media and other collaborative platforms within the public sector have resulted in a dynamic recordkeeping environment that demands a new approach. In short, a very significant portion of data is now created and stored online—and government organizations need to quickly adapt to this reality.

This paper will discuss the factors responsible for the recent evolution of public-sector recordkeeping and examine the characteristics of modern digital records that make them so challenging to deal with. It will do so by making use of concrete examples like website content, social media accounts, instant messaging services, and team collaboration tools. Lastly, the paper will conclude by offering an example of a federal agency that has implemented successful strategies for dealing with modern online records.

The Implications of Directive M-19-2

Undoubtedly, one of the biggest drivers of current recordkeeping changes within federal agencies is [Directive M-19-21](#). Originally issued by NARA on June 28, 2019, the final deadline for compliance has arrived, which means agencies must now:

- Manage all permanent electronic records in an electronic format
- Create, manage, and maintain permanent federal records—and their metadata—in an electronic format
- Have closed their individual records storage facilities and transferred any inactive or temporary paper records to a Federal Records Center or Commercial Storage Facility
- Maintain a records management program that complies with the Federal Records Act

What does this mean for federal agencies in practical terms? According to [the organization Collabware](#), Directive M-19-21 demands that organizations do the following:

- Develop a clear digitization strategy for the conversion of all permanent records to digital formats
- Put automated systems in place to capture, categorize, and retain all electronic content (with metadata)
- Develop measures to ensure the integrity and preservation of all electronic records throughout the entire data lifecycle
- Develop integrated management capabilities for electronic records, regardless of where any particular records are being stored
- Develop concrete methods for quick and accurate responses to FOIA requests
- Perform regular program audits to validate the effectiveness of records management programs and maintain accurate records inventories
- Put policies and systems in place to facilitate the easy transfer of electronic records to NARA
- Align their recordkeeping policies with [NARA's ever-expanding definition of data types](#) that qualify as federal records

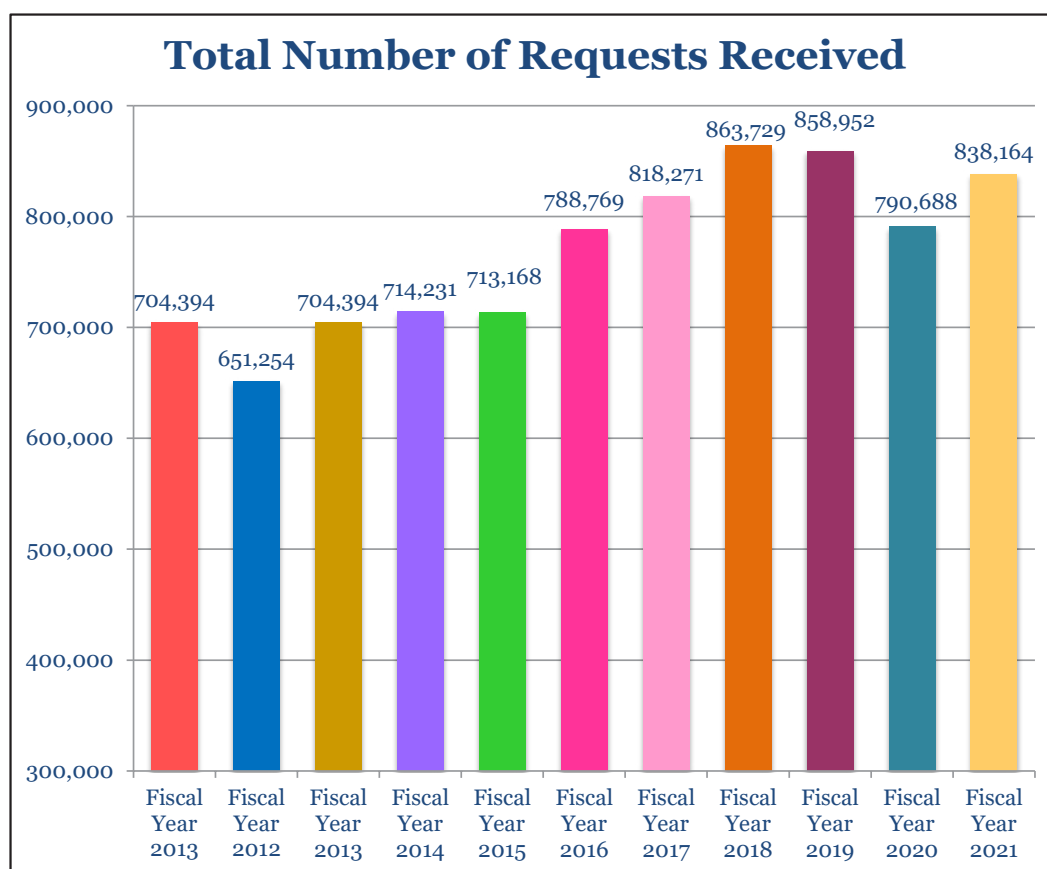
Given the volume, velocity, and variety of content, the complexity of websites, and the number of official social media accounts created and maintained by a federal agency, it quickly becomes clear how complying with these demands can be challenging:

- How can automated systems be put in place to capture, categorize, and retain all historical website and social media content?
- How does an organization develop measures to ensure the integrity and preservation of website and social media records when this data is being hosted by a third party such as WordPress, Facebook (Meta), or Twitter?
- How can this data be transformed and exported for easy transfer to NARA?

Given the cloud-based nature of this data, traditional tools and workflows do not work. What is needed is an approach that is geared specifically towards online data, is scalable, and does not require manual processes.

Dealing with Increased FOIA Requests

Federal agencies deal with an inordinate amount of FOIA requests every year—and there has been a general increase in requests over the last decade.



Source: [Summary of Annual FOIA Reports for Fiscal Year 2021](#)

Of course, the numbers above do not offer insight into the true scale of FOIA requests, since they don't reveal the size of the particular requests. Some requests are simple, asking for only a handful of documents. Others, however, can demand thousands upon thousands of files that need to be found, exported, and sometimes even redacted.

The Office of Information Policy's (OIP's) [Summary of Annual FOIA Reports for Fiscal Year 2021](#) reveals that:

- Federal agencies had 153,227 backlogged requests at the end of 2021—the highest since 2014
- The average time to process simple requests was 32.99 days
- Complex requests were typically processed in 40 days or fewer—but 15% of requests took between 100 and 400 days to process
- Agencies received 15,468 administrative appeals and processed 15,522
- 5,363 full-time FOIA staff were devoted to the processing of requests in 2021
- More than \$561 million was spent on FOIA activities in 2021—93% of the total costs was attributed to the administrative processing of requests and appeals, while 7% was attributed to litigation-related activities

Again, the shift to cloud-based applications and storage solutions require a rethinking and expansion of traditional FOIA processes. As organizations make increasing use of online tools to communicate and collaborate, the nature of many FOIA requests are also changing.

As one example, many agencies made use of their official websites and social media channels to update the public on the latest COVID information in the midst of the pandemic. And given the unprecedented nature of the event, this information changed regularly, often on a daily (or even hourly) basis. As a result, there have been a lot of FOIA requests for specific information that appeared on a website or social media account on a particular day. Fulfilling these requests has been difficult. Not only does this data need to be found and exported in a format appropriate for a FOIA request, but if it has been edited or deleted on the original platform, it can be almost impossible to retrieve.

In the final section of this paper, we will discuss a federal agency that dealt with precisely this challenge of processing FOIA requests related to COVID information on its website.



SECTION 3

The Need for Transparency

At its core, the Freedom of Information Act is aimed at improving government transparency. For this reason, subsection (a)(2) of the act demands proactive disclosures. Agencies are expected to proactively make non-exempt information from certain categories of records available to the public without waiting for specific requests to be received. These categories are:

- Final agency opinions and orders rendered in the adjudication of cases
- Specific policy statements that are not published in the Federal Register
- Administrative staff manuals and instructions to staff that affect members of the public
- Records that have become or are likely to become the subject of subsequent requests or those records that have been requested three or more times—referred to as “frequently requested records”

Beyond reducing the number of FOIA requests that must be processed, the proactive disclosure of the above categories of documents (especially frequently requested records) increases government transparency and builds public trust. Consequently, many agencies are opting to take a best-practice approach and make as much information as possible publicly available—including historical website and social media content. The final section of this paper will provide an example.

The Challenges of Online Data

When it comes to meeting the demands of the Freedom of Information Act, Directive M-19-21, and other recordkeeping requirements, online data (such as website content, social media accounts, text messages, and team collaboration platforms) introduce unique challenges. Some have already been mentioned briefly, but below is a list of some of the most common challenges that agencies face:

- **Taking ownership of online data:** Because data is cloud-based and owned by third parties, taking control of this data and managing it in a way that aligns with NARA and FOIA requirements is difficult.
- **Reliance on manual processes:** In order to meet recordkeeping requirements related to online data, agencies often resort to screenshots and other manual methods that are incredibly time-consuming and do not provide the metadata needed to be truly compliant.
- **Capturing content regularly and at scale:** Related to the point above, organizations struggle to capture content on a frequent basis when dealing with massive websites, active social media accounts, etc., which results in significant record gaps.
- **Capturing complex and dynamic content:** Data sources like websites and social media posts contain complex content, like Java, media files, links, etc., that is difficult to capture reliably and in their entirety. Content can also be edited or deleted, creating multiple records of a single post or webpage.
- **Reviewing and reproducing captured content:** Due to the complex nature of online content, reviewing and reproducing this data with an accuracy sufficient to meet FOIA and NARA requirements is challenging.

To see how these challenges can be overcome, the next section will examine how the U.S. Department of Health & Human Services (HHS) deals with its massive volume of website records.

Recordkeeping Example: U.S. Department of Health & Human Services

According to the Office of Information Policy's (OIP's) [Summary of Annual FOIA Reports for Fiscal Year 2021](#), a mere five federal agencies received more than 75% of all FOIA requests in 2021. These agencies are:

- Department of Homeland Security (DHS)
- Department of Justice (DOJ)
- Department of Defense (DOD)
- Department of Veterans Affairs (VA)
- Department of Health & Human Services (HHS)

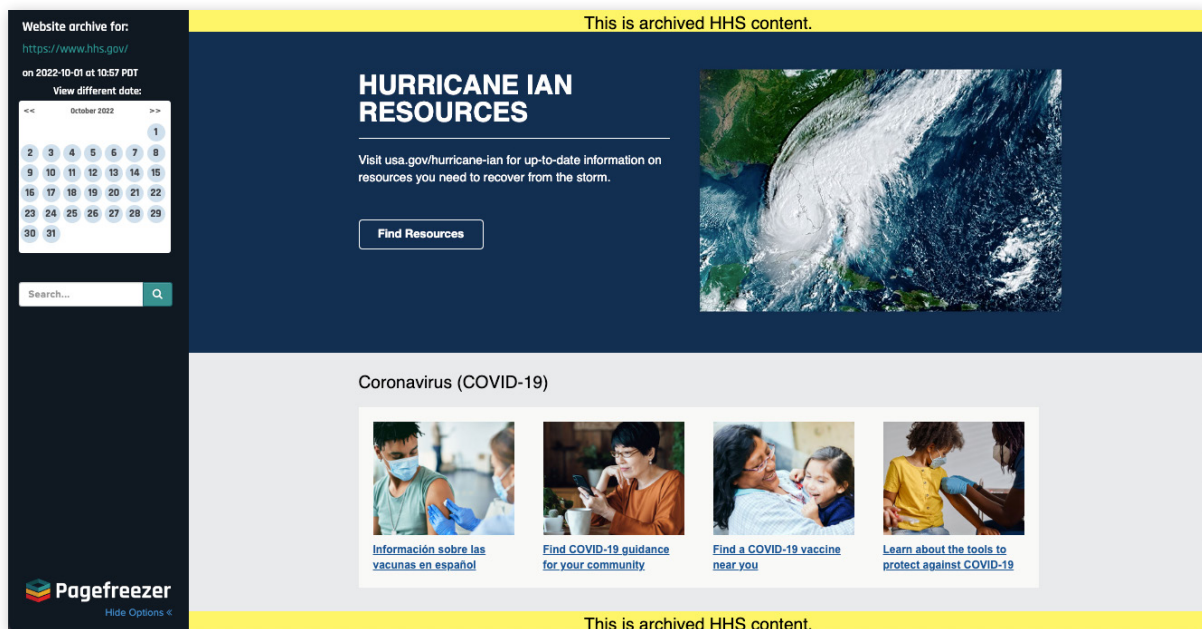
In this section, we will look at the Department of Health & Human Services, an organization that saw a particular uptick in FOIA requests during the COVID-19 pandemic.

Due to the dynamic and uncertain nature of the pandemic, information on official HHS websites needed to be updated frequently—often on a daily basis. So, to improve transparency and reduce FOIA requests related to previous website content, the agency launched an initiative to identify and implement a solution that would allow it to easily share historical web content with the public.

After evaluating its options, HHS decided to implement [Pagefreezer's web archiving solution](#). Pagefreezer allowed the organization's archiving efforts to scale across its entire public web footprint, automating the daily capture of more than one million webpages across numerous domains.

This captured data was then made available via Pagefreezer's browser-based dashboard, allowing hundreds of users and stakeholders across the organization to quickly search, discover, and process web records.

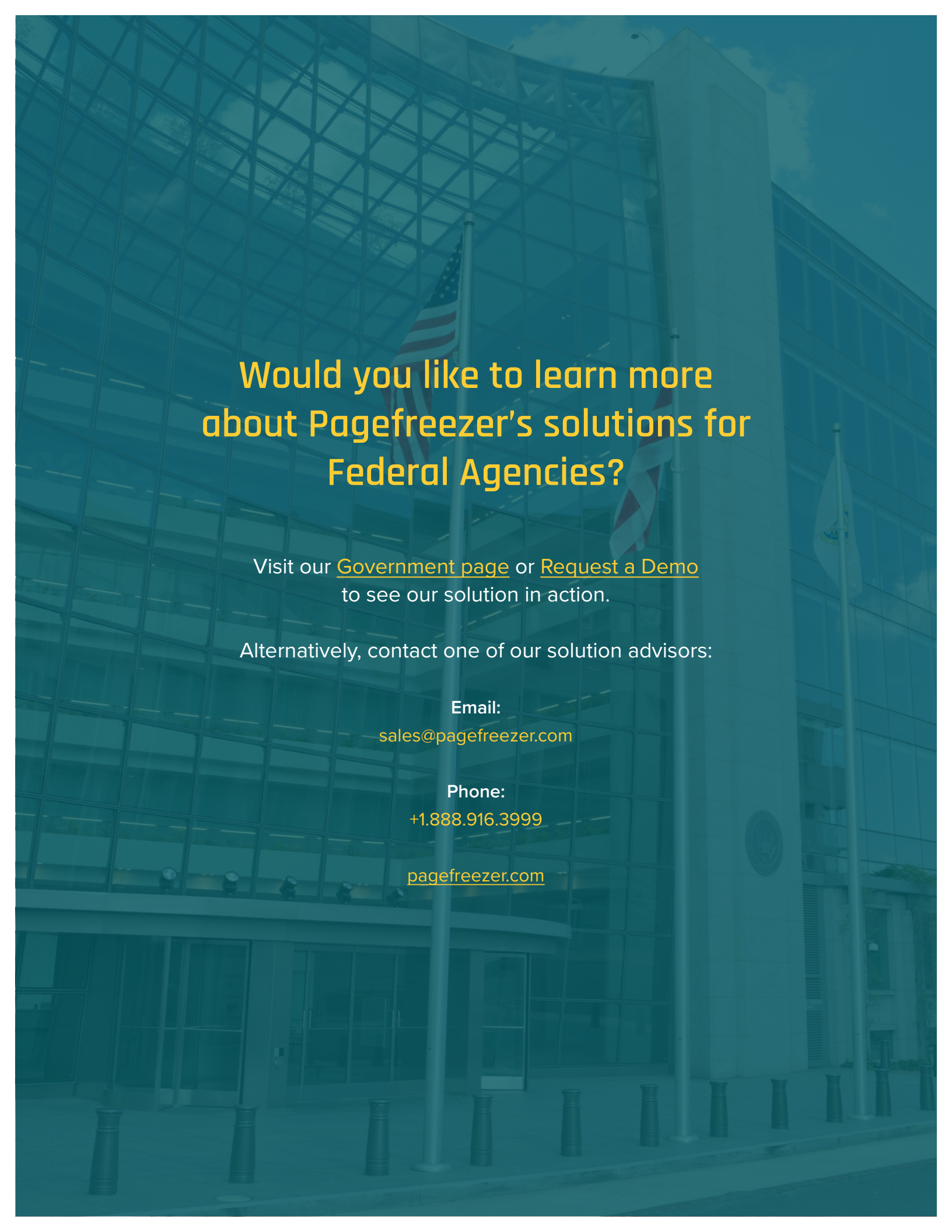
Moreover, the Department of Health & Human Services opted for the creation of a [public emulation portal](#) that would allow members of the public to view any historical website content, simply by visiting the portal and selecting the date they're interested in.



Members of the public can access the US Department of Health & Human Services to view and interact with historical website content.

By implementing this archiving solution, the HHS was able to:

- Reduce the number of FOIA requests from journalists, special interest groups, and general members of the public
- Improve transparency by making historical website content instantly accessible through its public archive portal
- Reduce the amount of time and resources that had to be allocated to the fulfillment of FOIA requests
- Streamline the processing of website records for use in eDiscovery, compliance audits, records management, etc.



Would you like to learn more about Pagefreezer's solutions for Federal Agencies?

Visit our [Government page](#) or [Request a Demo](#)
to see our solution in action.

Alternatively, contact one of our solution advisors:

Email:

sales@pagefreezer.com

Phone:

+1.888.916.3999

pagefreezer.com