



Making the Case for PDF/A

Document Processing &
Archiving in the Digital Age



Table of Contents

1. Making The Case For PDF/A	3
2. Archiving & Processing Documents In The Digital Age	4
3. The Challenges And Complexity Of Document Archiving	5
4. Converting Documents To A Standard Electronic Format	6
5. PDF Evolves Into The Electronic Document Standard PDF.....	8
6. PDF As A Records Management Document Solution	10
7. PDF/A: Document Solution For Archiving And RM	12

Making the Case for PDF/A

Digital media needs to be archived, searched, and analyzed – now more than ever. PDF is the only standard that has the flexibility, industry trust, and the International Organization of Standardization (ISO) certification to be used as an electronic document archiving standard.

PDF/A or Portable Document Format Archival, is the version of PDF specifically designed to provide the set of features required to encapsulate all of a document's properties, with no information loss. It is designed to ensure that the content of a file can be preserved for decades. With PDF/A, one can reliably replicate the exact look and feel of a document years later. As an archive file, it has been designed to prohibit modifications because of legal and historic preservation issues.

Tools like Foxit PDF Editor supports converting general files to PDF/A and allows you to work with PDF/A-compliant documents with ease.



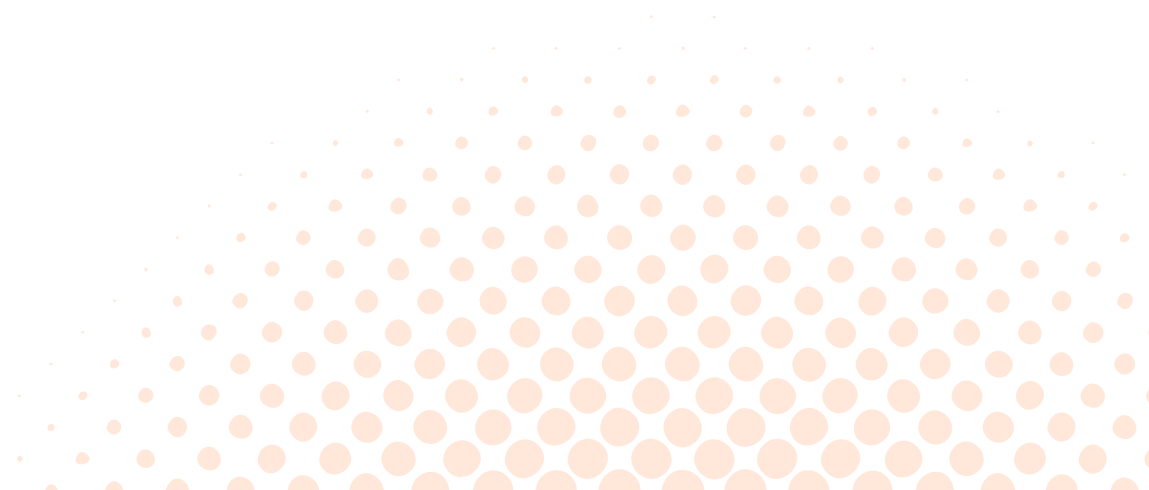
Archiving & Processing Documents In The Digital Age

In this paper, we motivate the logical imperatives that have made PDF the de facto international standard for document processing and archiving. We further describe how PDF/A, a variation of PDF that is specifically designed for long-term document archiving, is the next step in electronic document archives. Companies and government agencies that take archiving and records management seriously are increasingly requiring that PDF/A be the format of choice for their documents.

In the legal field, there is a growing acceptance and requirement for standardized electronic filing formats in both civil and criminal cases. Courts are now shifting towards the use of PDF/A for electronic submissions to ensure the long-term accessibility and compliance with records management. This is driven by the need for documents to be viewable and searchable and subsequently addressing concerns related to document integrity and authenticity.

Many jurisdictions are now adopting rules that specify PDF/A as the acceptable format for documents. This shift emphasizes the importance of preserving documents and their security features, such as digital signatures. Such legal requirements aim to facilitate smoother case management, improve document retrieval processes, and ensure that electronic files and records remain reliable and usable even after decades.

In this whitepaper, we also explore some of the challenges related to document archiving. These challenges, though not always the same, apply to physical, captured image, and “digitally born” files. We will also discuss how and why the industry has gravitated towards relying primarily on PDF for document management, workflow processing, and archiving. Finally, we will dive into the next natural progression: the adoption of PDF/A as the standard for records management and long-term document archiving.



The Challenges And Complexity Of Document Archiving

Document records have been a crucial part of business, commerce, and government for thousands of years. The Dead Sea Scrolls serve as a prime example of effective document preservation, maintained over centuries by applying appropriate archival techniques (e.g., tightly wrapped, with no exposure to light or moisture). Until very recently, documents were easier to define and archive. Historically, they were written on parchment or paper, which were stored in controlled environments for preservation.

Today, documents are primarily paper and electronic, and as their complexity increases, organizations need effective, reliable policies for archiving them. These must include effective and reliable policies for metadata, hyperlinks, and executable code, such as JavaScript, to support modern document functionalities.

As digital data increases, there is significantly more information for companies and organizations to manage. Common document types include attachments, incoming surfacemail, and faxes, which often encompass forms, invoices, checks, and contracts. Archiving a document should maintain the original look and feel, including margins and pagination. Formats like Word and other Microsoft Office documents, which depend on platform settings and viewer preferences, are unsuitable for archival purposes. Image formats such as JPEG and TIFF may preserve the appearance of the original document but often lose essential functionality like text searching and metadata.

In the legal field, the importance of effective document archiving cannot be overemphasized. Legal professionals must ensure that documents are not only preserved but also retrievable and searchable, especially for compliance and

discovery purposes. In the past, retrieving historical documents often involved searching through microfiche reels, which recreated newspaper pages faithfully but were time-consuming and inefficient for search queries. Additionally, microfiche lacked effective metadata insertion, making it difficult to locate specific information.

The shift towards electronic filing and archiving in courts has underscored the necessity for standards like PDF/A. This format ensures that legal documents remain accessible and retain their integrity over time, supporting the rigorous requirements of legal proceedings. As case law increasingly relies on digital records, the challenge remains to implement comprehensive policies that address both the complexity of modern documents and the legal standards for document management.



Converting Documents to a Standard Electronic Format

An increasing number of legal firms have compelling interest in converting their incoming documents into a single electronic format. There are several reasons for this. First, when all documents are combined in the same file, it is much easier to search, view, and print them. This file format needs to be as functional as possible, while ensuring that the integrity of the original document is preserved. Having one standard format simplifies document workflows. If all files are converted to one format, it allows for easier viewing, classification, routing, data extraction, analysis, and commenting, all using the same software.

Second, archiving “digitally-born” or native electronic documents, has its own challenges similar to archiving paper and captured paper documents. With respect to records management and archiving of electronic files, a native approach would retain its native electronic format and import them to the company’s record management system. This method may be simple and straightforward, but it can be risky. Retaining documents in their native format is not aligned with the best practices for document archiving and records management as suggested by National Archives and Records Administration (NARA) and Association of Records Managers and Administrators (ARMA).



Electronic Document Archives Preserved In Their Original Format.

Keeping corporate documents in their native electronic format can lead to significant issues. These problems include:

- 1. Viewing Challenges:** Without a policy to convert documents to standardized file format will require supporting different viewer tools.
- 2. Inconsistent Display:** A document produced in one office location, such as an MS Word file generated in a company's Mexico City office, may not display correctly in the same company's New York office. This inconsistency can arise because the Windows system fonts may differ between the locations.
- 3. Inconsistent Appearance:** There is no guarantee that an electronic document will maintain its look and feel when generated on one machine and then displayed on another.
- 4. Margin and Pagination Changes:** Variations in margins and pagination can lead to incorrect page references.
- 5. Legal Document Retention:** It is hard to show that a given original electronic file has not been modified. Using the PDF format, this could be established with a digital signature.

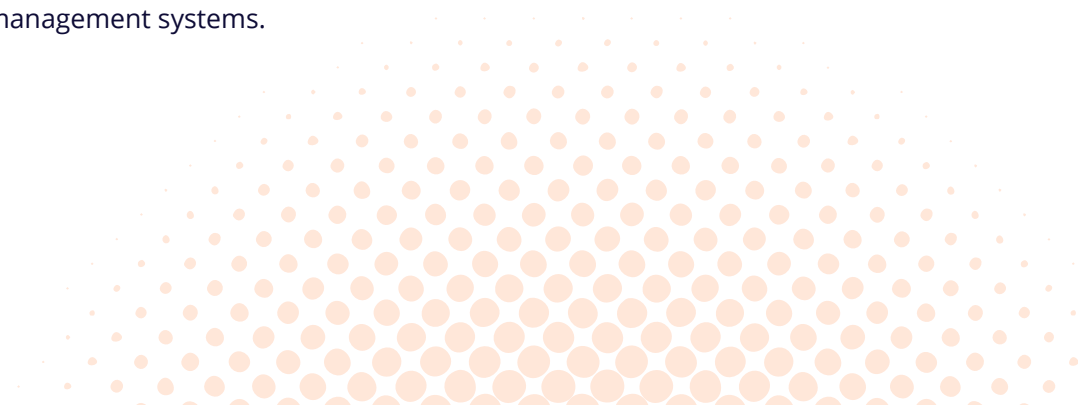
These are some of the reasons that archiving documents in their native format is often inadequate, both from a legal and regulatory perspective.

Electronic Documents Preserved In Image Format

Archiving documents to image formats, such as TIFF or JPEG, can present several challenges, especially in the legal field. For the past two decades, converting documents to TIFF format has been used extensively for processing and archiving corporate and government records. This electronic imaging (capture) was actively used in lieu of original files, which were either put into "deep storage" or discarded, reducing the need for microfiche.

Some of the advantages of converting documents to TIFF format are that it retains the original page layout, which is important for legal documents that require precise formatting. It cannot be easily modified, helping to ensure the integrity of important legal documents. It can also be widely viewed and printed, making it suitable for sharing among legal teams and clients. Despite some significant limitations, many organizations still used it as a document archiving format.

However, as new trends are now being set to improve the management and archiving documents in the legal industry such practice may face challenges. TIFF files are not searchable, which can make specific information retrieval difficult during document reviews or litigation. TIFF files may also not comply with the modern requirements for e-discovery documents, which tend to favor formats that allow for easier analysis and integration with document management systems.



PDF Evolves into the Electronic Document Standard PDF

In the last decade, with the increasing number of both captured paper and “born digital” documents being generated, PDF has become the preferred choice for many organizations, especially in the legal industry, for handling their document workflow and retention needs. There are several reasons for the increasing popularity of PDF format in the US and globally.

These include:

- 1. Widespread Adoption:** PDF has become the standard format for both electronic and scanned documents, making it the optimal choice for legal filings and documentation.
- 2. Accessibility:** PDF readers, such as Foxit PDF Reader, are available for viewing and editing PDF files, ensuring that legal professionals can easily access and collaborate on important documents with ease.
- 3. Advanced Features:** PDF supports features such as smart compression and digital signatures, which are essential for maintaining document integrity and security.
- 4. Efficient Workflows:** PDFs are effective in advanced workflows, including automated document classification and routing through the use of inserted metadata, which streamlines processes such as e-discovery.
- 5. Support for Scanned Documents:** PDF can reliably support captured or scanned documents using the image with hidden text format, allowing for easier searching and retrieval of information.

6. PDF/A for Archiving: PDF/A is a specialized version of PDF designed for long-term document archiving, ensuring compliance with legal retention requirements.

Unlike many electronic formats such as Word documents or Excel spreadsheets, PDF is an ISO standard, meaning its specifications are published and transparent, allowing third-party software to both read and write PDF files. While some ISO standards may go decades without updates and formats like TIFF and JPEG have remained largely unchanged for over 20 years, the PDF standard has seen frequent revisions, with more than 10 updates since its initial standardization. PDF was formalized as an ISO standard in 2008 and continues to be enhanced regularly.

This regular innovation gives PDF a significant advantage. The advanced compression technologies, including JBIG2 (bitonal captured files) and JPEG2000 (grayscale and color files) have been integrated into PDF specifications. Additionally, features like document encryption and digital signatures are fully supported.

Overall, PDF formatting gives the user all the benefits of a proprietary format, with regular updates and incorporation of the latest technologies, coupled with the reliability and transparency of an open ISO standard. This makes PDF an ideal choice in the legal industry where security, accessibility, and efficient document management solutions are a must.

Evolution of Document Capture: A Factor in the Increased Use of PDF

In this section, we'll discuss the evolution of document capture and its relationship to PDF format becoming the primary document standard. Over the last decade, there has been strong growth in the use of document capture technology. Whereas, more than a decade ago, most corporate scanning was highly centralized. Today, almost every modern office is equipped with a multi-functional printer (MFP) that can digitize documents and store or send them in an email. This shift, along with efficiency and cost savings of maintaining electronic archives instead of physical ones, has driven the widespread adoption of document capture.

The return on investment (ROI) for document capture is straightforward: instead of having two distinct workflows, integrating workflows can eliminate the need for separate processes for paper-based and electronic documents. Businesses or organizations can convert all incoming paper documents into electronic files through scanning, streamlining operations into a single electronic-based workflow.

This evolution from paper-based document workflows to primarily electronic document workflows, which involves capturing incoming paper documents, has significantly contributed to the popularity of the PDF format. For instance, in a law firm managing client intake and case documentation, one-third of client

forms and evidence submissions might be received digitally in PDF format, while the remaining two-thirds arrive via email attachments or fax. By converting all paper and electronic documents into a single, standardized workflow, PDF becomes the logical choice. Electronic submissions are often already in PDF, while mailed or faxed documents can be easily converted into searchable PDF images with hidden text. This standardization ensures consistent formatting, easier document indexing, and more efficient case management, making PDF the preferred format for legal professionals handling sensitive and complex information.

As we move swiftly into the digital age, it is inevitable that electronic archives will surpass paper-based archives. The advantage of using PDF as the standard for electronic archives is that it has the benefits of being able to rapidly evolve and incorporate the latest technological advances, while still maintaining the ISO certification as an accepted, international standard. This adaptability makes PDF an ideal choice by legal professionals who consider security, efficiency, and compliant document management solutions necessary to keep up with digital advancements.



PDF as a Records Management Document Solution

The widespread adoption of PDF as the leading format for document workflows is significant. However, this alone does not guarantee its suitability for legally reliable use in regulated fields such as records management (RM). Corporations and government agencies often enforce strict RM and document retention policies, guided by standards like ARMA International's Generally Accepted Principles of Record keeping. These principles require that records be:

- **Authentic**
- **Reliable**
- **Complete**
- **Unaltered**
- **Usable**

Fortunately, PDF solutions like Foxit effectively support each of these objectives, making it a strong choice for records management and archiving in compliance-driven environments.



Authenticity:

To verify that a record is genuine and created by a specific individual at a particular time, metadata plays a crucial role. Metadata is essential for long-term preservation, as it allows key details—such as author, creation date, subject, and keywords—to be embedded within the document. This practice enhances document portability since information is stored both at the database level and within the file itself. In a controlled document workflow, the creation, receipt, and transmission of records can be systematically managed, ensuring that each record creator is properly authorized and identifiable. Tools like Foxit software can programmatically insert metadata into PDF files and apply digital signatures to confirm authenticity and prevent tampering.

Reliability:

A reliable record must accurately reflect the event or transaction it represents. To achieve this, documents should be created promptly by individuals with direct knowledge or through automated processes regularly used for such tasks. PDF standards support the design of automated systems that generate digital PDFs directly from source applications, ensuring a consistent and reliable output for records management purposes.

Integrity:

Preserving the integrity of records means ensuring they remain complete and unaltered. Unauthorized changes must be prevented, while any approved annotations, additions, or deletions should be carefully tracked. PDF files can be secured with encryption and password protection, restricting access to authorized users only. Granular permissions can control viewing, editing, and printing rights. Additionally, digital signatures can lock the document from further modifications once signed, further safeguarding document integrity.

Usability:

Records must be easy to locate, retrieve, render, and understand. PDF excels in this area through consistent full-text search capabilities. Born-digital PDFs are text-searchable from creation, while scanned documents can be made searchable using Optical Character Recognition (OCR). Furthermore, PDF's ability to encapsulate XML metadata ensures long-term usability, preserving critical record details for future reference.

PDF's combination of authenticity, reliability, integrity, and usability makes it a trusted format for the most demanding corporate and government records management (RM) applications. As industries, including the legal sector, continue to embrace digital transformation, the versatility, security, and compliance capabilities of PDF make it an essential tool for modern document workflows.

PDF/A: Document Solution for Archiving and RM

To address the need for a standardized electronic format for document preservation, the International Organization for Standardization (ISO) introduced PDF/A as the preferred format for long-term document archiving. This initiative began in 2002 and was officially approved in June 2005. The development of PDF/A was a joint effort led by the Association for Information and Image Management (AIIM) and the Association for Suppliers of Printing, Publishing, and Converting Technologies (NPES), with contributions from a diverse group of librarians, archivists, software developers, imaging experts, government agencies, and industry professionals.

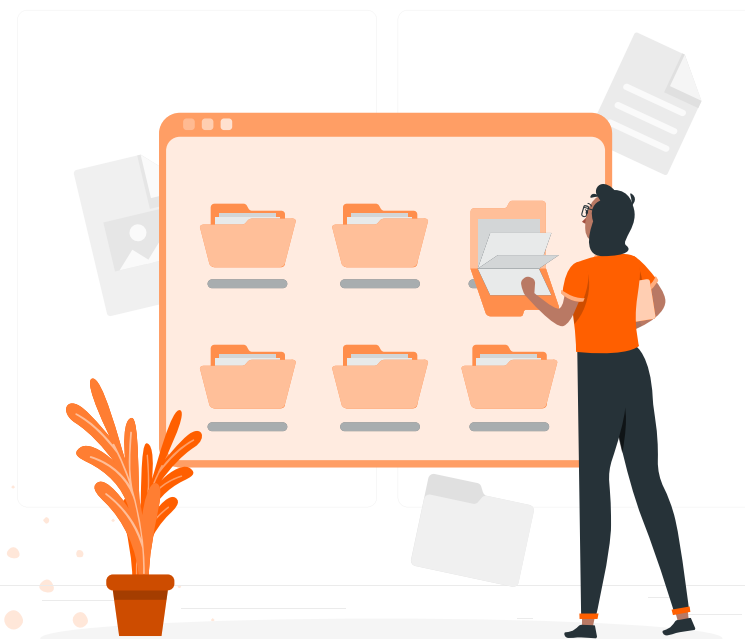
According to ISO 19005-1, PDF/A “provides a mechanism for representing electronic documents in a manner that preserves their visual appearance over time, independent of the tools and systems used for creating, storing, or rendering the files.”

PDF/A was specifically designed to ensure long-term accessibility by modifying the standard PDF format. While standard PDFs support a wide range of features, such as multimedia and encryption, these elements can make documents difficult to preserve as technologies evolve. Additionally, traditional PDFs may depend on external content or proprietary features, which could render parts of the document inaccessible in the future.

PDF/A ensures long-term usability by requiring documents to be fully self-contained and omitting features that could hinder preservation.

PDF/A maintains the familiar appearance and functionality of standard PDFs but restricts certain elements that are incompatible with long-term archiving, including:

- **Encryption**
- **Embedded files**
- **External content references**
- **Transparency effects**
- **Multimedia elements**
- **JavaScript**



While these restrictions make PDF/A highly suitable for secure archiving, they can present challenges in industries requiring strict document access controls. For example, financial services and investment banking often need to restrict viewing or printing of sensitive records. Though standard PDF encryption can address these needs, such protections are not supported under PDF/A specifications due to its focus on long-term accessibility.

PDF/A remains a reliable and trusted solution for organizations seeking to preserve documents with authenticity, integrity, and usability for years to come. PDF/A will maintain the static visual appearance of electronic documents over time, while maintaining the capacity for metadata. Metadata is essential in long-term preservation because it allows users to insert identifiable information regarding the details of a document such as author, date, subject, keywords, and more. This metadata enhances the portability of a database by ensuring that information is preserved both at the database level within the document itself.

PDF/A has proven to be an essential standard for the legal sector, ensuring the long-term preservation, authenticity, and accessibility of critical documents in compliance-driven environments. Its ability to safeguard document integrity while maintaining usability makes it ideal for archiving case files, contracts, evidence, and other sensitive legal records. For firms seeking a reliable solution to meet these rigorous standards, Foxit PDF Editor provides a comprehensive platform for creating, converting, and verifying PDF/A documents. With advanced tools for metadata management, digital signatures, and secure access controls, Foxit empowers legal professionals to streamline document management, ensure compliance, and preserve the integrity of their records for years to come.





FOXIT | FEBUARY 2025