

TOP 10 WAYS

PDF HELPS YOU TAKE CONTROL OF DIGITAL DOCUMENTS

What can one do with the PDF format, and where is it going? For two decades, PDF has been the global standard for document viewing and cooperation. But despite this, many of the format's features remain largely unknown even to many IT professionals. PDF is continuously evolving, and its current advantages and flexibility may surprise you. The following provides an overview of the PDF's current assets.

PDF is a file format that excels in the face of all major challenges presented by the ongoing explosion of big data technology. While it is largely known for its strengths in terms of viewing, sharing, archiving, and signing, the format also shines in terms of programmability. PDF displays tremendous potential in terms of automating transactions and improving flows of collaboration because it is far more than the digital copy of a paper document. Metadata, structure, forms, and other file type attachments packaged together, turn it into an extremely versatile envelope for your content.

The strengths of the PDF format make it difficult to imagine a well-designed document solution that does not integrate PDF files in a central way. The format has a great deal to offer because it responds by design to most of the challenges encountered in building a digital document workflow, whether they are related to data processing, long-term archiving, digital signatures and security, or cross-platform sharing.

THE EXPLOSION OF DOCUMENTS AND TECHNOLOGIES

Our economy runs on documents: there are entire sectors of the economy – IT, finance, accounting, sales, marketing, legal services and many more – where treatment of information represents all of the business. In these sectors, much of the difference between runaway success and rapid failure is made by decisions on document management. Documents are business-critical, and their number and complexity increases every year by orders of magnitude.

	Customer communications
	Accepting and confirming orders (receipts, e-tickets)
	Internal communications (banking records, business plans, contact information)
	Legal agreements (contracts, approvals, insurance policies, leases, property deeds)
	Reporting output (financial statements, tax returns)
	Product design (engineering diagrams)
	Software licenses and product keys
	Customer relationship management data
	Payroll records

Innovation is changing the nature of documents...



Your mobile app

collects customer data via form and flattens it into a document



Your accounting system

produces personalized invoices to be sent to customers



Your software

seamlessly generates customer orders from a database



Your e-commerce

web portal produces barcoded e-tickets

Advancements in browser, HTML, CSS and mobile technology have enabled us to introduce interactive and richly formatted publication experiences to all. Even more, branded and customized applications produce new types of interactive documents for all devices, forms and functions.

... while legislation is keeping them in place

Because any file can be conceivably used in court, organizations must be able to prove the integrity and authenticity of the files at their disposal. As a result, these interactive documents need to be reconciled by law with document management retention and destruction practices and standards. Unfortunately, most modern IT systems are challenged to continuously innovate while maintaining complete reliability.

The goal has shifted over the past two decades from merely generating, printing and transporting documents to managing big data intelligently.

Professional IT systems are supposed to increase productivity and improve business processes. However, most modern IT systems are challenged to continuously innovate. The numbers below illustrate how elusive the goals of the information revolution remain. Still true today, most businesses are bogged down by sub-optimal search processes even with common practices, such as customer correspondence.



The cost to business can be overwhelming:

- A great deal of company information is vulnerable to theft or loss, either because it exists only as paper documents or in unencrypted digital files that are not backed up
- Filing, organizing, and retrieving documents – physical or digital – takes too much time
- Documents and files often get lost or misplaced, costing valuable time to find or replace (if replacement is an option)

Ideally, we would have a single file format that is nimble enough for the newest technology but embeddable within legacy business systems to mitigate risk

WHY PDF

PDF (the Portable Document Format) was first introduced in the early 1990s as a way for documents to be shared and displayed on any computer using any operating system. Two and a half decades later, PDF has evolved into a versatile format for addressing our contemporary business challenges caused by the explosion of digital data.

- Self-contained, totally portable across platforms
- Tremendous flexibility and power in representing documents and forms
- Easy to create using hundreds or thousands of independent implementations
- May be secured for distribution and protection
- May be authenticated with digital signatures
- May include semantics to assist in content reuse and accessibility
- The non-proprietary, democratically-managed, globally-accepted de facto standard for electronic documents

What can PDF contain?

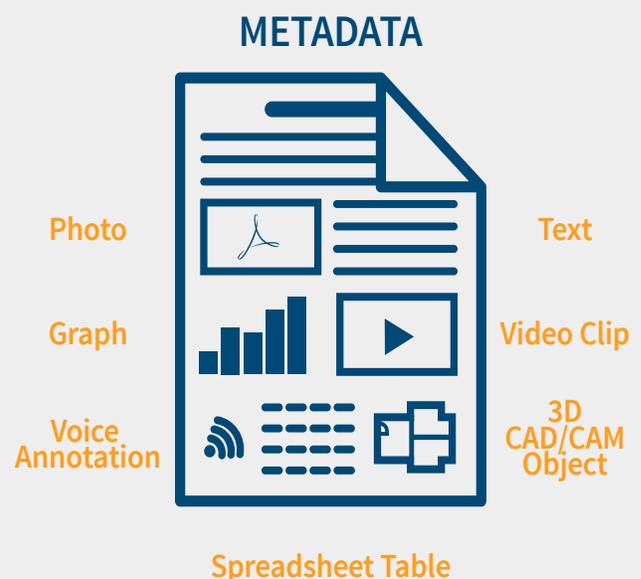
PDF is a “self-contained” document. It can contain content, other file types and interactive elements, as well as information about these elements:

Content

- Textual content: words, sentences, numeric values, tables
- Line art: curves, lines, shapes
- Pictures: raster images
- Multimedia and interactivity
- Markup and comments
- Display and navigation instructions
- Attachments and embedded files

Information about the content

- Semantic meaning and alternate information
- Notes and annotations
- Metadata & properties
- Measurements and geospatial information
- Attachments, including source documents



PDF: AN OPEN STANDARD

PDF has been an open, international standard since 2008 (ISO 32000) and is therefore freely implementable. Many software tools, either commercial or free, are available for PDF treatment.

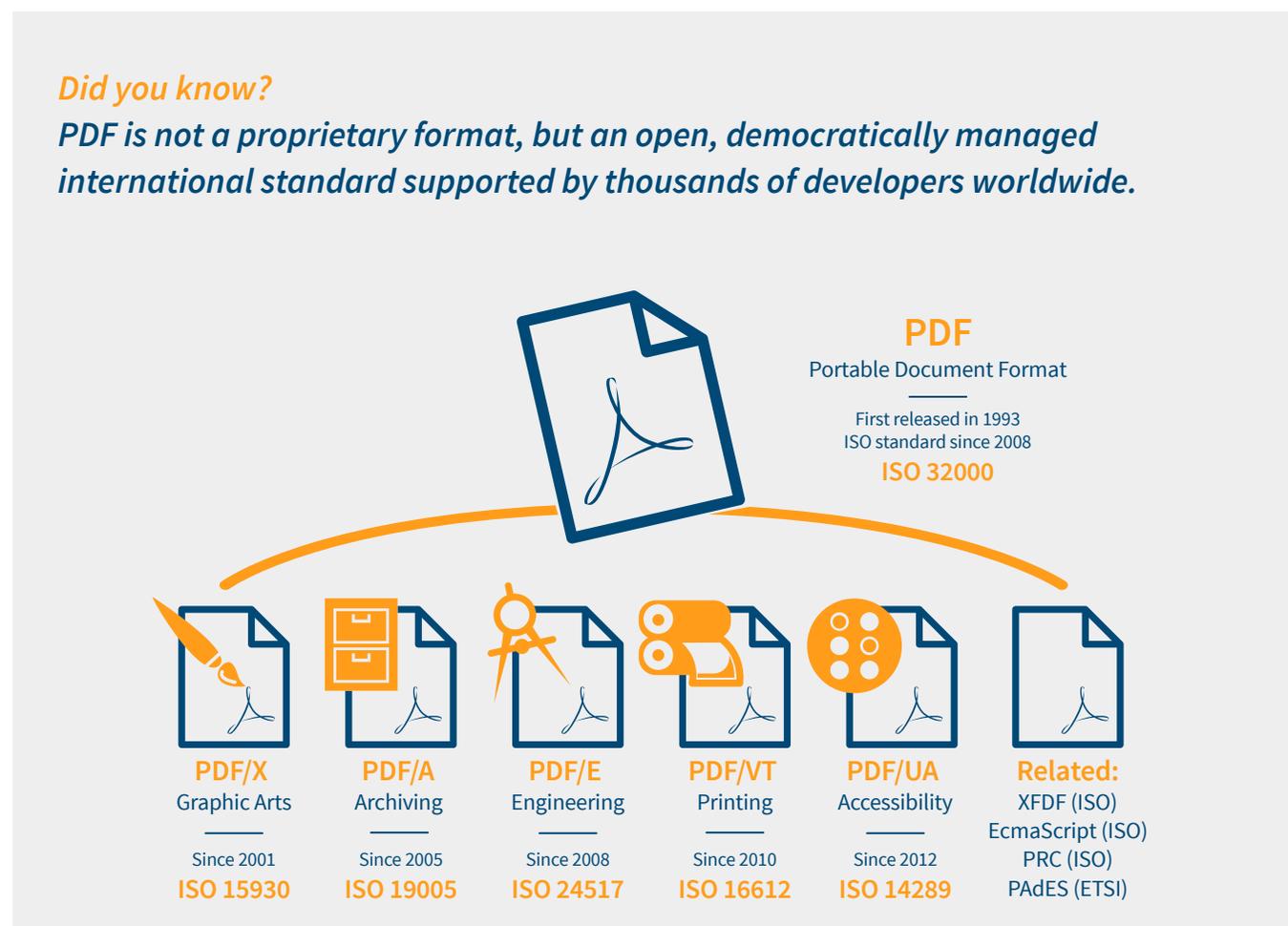
Different styles of PDF exist, each one being tailored to specific business needs. Many industry standards, such as PAdES for digital signature, and ZUGFeRD for secure digital invoicing are also based on the major PDF standards.

Some of the most business-critical applications include:

PDF/A	Specialized for the digital preservation of electronic documents e.g. font embedding
PDF/UA	Designed to ensure accessibility for people with disabilities who use assistive technology
PAdES	Ensures that digitally-signed documents may be used or archived for many years
ZUGFeRD (PDF/A-3)	Allows the exchange of invoices between supplier and payer without prior arrangements

Did you know?

PDF is not a proprietary format, but an open, democratically managed international standard supported by thousands of developers worldwide.



MEETING BUSINESS CHALLENGES

PDF is much more than an “image” of a paper document. As we saw earlier, it can have **internal structure, interactivity, workflows, forms, and various file types**, that can be understood not only by humans, but also by machines. Such features make it a central file format to many web and in-house enterprise systems, with the potential of:

- **Automating transactions** for thousands of documents
- **Personalizing communications** with linking to data sources
- **Speeding up reviews and approvals** with digital signature
- **Managing Big Data** by structuring and indexing files
- **Archiving and storing** for the long-term with embedded fonts and other file types
- **E-discovery and legal compliance**

...and so much more. For most document management challenges, PDF is a reliable backbone.

The following Top 10 unique applications of PDF make it an indispensable tool for any business or government looking to mitigate risk, or get a handle on the digital document chaos.

PDF CAPABILITIES	BUSINESS BENEFITS
<p>1 LONG-TERM ARCHIVING PDF/A The PDF/A family of standards defines how to create electronic documents that remain consistently reproducible for several decades; PDF/A is by far the top choice for archiving by all large companies, governments or universities.</p>	<p>BUSINESS CONTINUITY PDF largely solves the problem of technology obsolescence. Furthermore, PDF usage makes the containment of IT problems easy when they occur because PDF files are autonomous blocks of information, independent from each other.</p>
<p>2 DATA EXTRACTION AND ACCESSIBILITY Logical structure and PDF/UA A “tagged” PDF includes document structure and semantics information to enable reliable text extraction and accessibility. Tagged PDF defines a set of standard structure types and attributes that allow page content (text, graphics, and images) to be extracted and reused for other purposes. PDF/UA ensures screen readers, screen magnifiers, joysticks and other technologies can navigate and read electronic content.</p>	<p>MANAGING BIG DATA Today, 90 percent of all current digital data is unstructured, i.e. from sources that are not in traditional databases, such as videos or images data. Accessibility offered by PDF allows companies to create smart documents that can be read not only by humans, but also by machines.</p>

PDF CAPABILITIES	BUSINESS BENEFITS
<p>3 DIGITAL SIGNING Embedded electronic signatures PDF has been explicitly compatible with embedded electronic signatures for a very long time. Digital signatures can provide a range of valuable capabilities, from tamper-protection to authentication and revocation.</p>	<p>FASTER APPROVAL PROCESSES AND SECURITY Nearly all countries are moving towards a legal framework that recognizes, under certain conditions, the same value to digital evidence as to paper documents. While not all governments are moving at the same pace, the PDF format is encouraged by many legal systems.</p>
<p>4 BUSINESS PROCESS MANAGEMENT Workflow programmability PDF allows the management of all workflow elements: data processing, role assignment, rule definitions etc. The common platform of PDF technology allows users to collaborate on document reviews online and offline.</p>	<p>STREAMLINED PROCESSES AND INCREASED COLLABORATION PDF programming is easily integrated with existing business applications (whether pertaining to chain management, ERP or payroll and human resources), resulting in greater productivity.</p>
<p>5 SECURITY AND PRIVACY Authentication and verification A PDF file may be encrypted for security using password or certificate encryption, and digitally signed for authentication. The digital signature function enables verification that the signed document has not been altered and was signed by someone the recipient trusts.</p>	<p>SECURE DOCUMENT EXCHANGE While any file format can be tampered with, PDF is most useful for creating and controlling feature-rich and secure digital signature workflows. It is compatible with industry best practices and is the basis for standardized document exchanges, such as invoices and contracts.</p>
<p>6 DOCUMENT STANDARDS International standard PDF is the de facto standard nearly everywhere, and increasingly de jure as well. Many governments require the saving of non-editable documents in PDF/A format (Denmark, France, Switzerland, Germany, and more). This has also been common practice in the U.S. for many years, especially in the court system.</p>	<p>LEGAL COMPLIANCE Current PDF versions adhere to all major international standards. In one move, the business manager can integrate his system within the international ecosystem of norms - examples include:</p> <ul style="list-style-type: none">• Section 508 (a U.S. federal amendment, which deals with issues of the access of the disabled)• PAdES (a norm of the European Telecommunications Standards Institute defining secure electronic signatures) and• ZUGFeRD (a German format for electronic invoicing with strong potential of becoming a European standard).

PDF CAPABILITIES	BUSINESS BENEFITS
<p>7 CROSS-PLATFORM SHARING PDF ubiquity The PDF format was originally created as an answer to the challenge of consolidating heterogeneous digital data over multiple platforms, and it remains the leader in the field. PDF viewers are ubiquitous on PCs and bundled on nearly all mobile devices.</p>	<p>SYSTEM INTEROPERABILITY Whether your applications are native, web, or hybrid, your business will have increased leverage over all platforms and formats. PDF guarantees near-universal accessibility, now and in the future.</p>
<p>8 INCORPORATING DATA SOURCES Dynamic and interactive forms The PDF format is widely used for submitting and importing data via interactive forms. The form fields' values may be submitted in a number of formats, including HTML, FDF and XFDF.</p>	<p>PERSONALIZED COMMUNICATION AND REPORTING PDF forms integrate seamlessly with any number of database solutions on the back end, resulting in quick processing of large quantities of data.</p>
<p>9 DOCUMENT FILING AND SEARCHING Descriptive metadata PDF files can contain a set of key/value fields, such as author, title, subject, creation and update dates. Also, using the Extensible Metadata Platform (XMP), PDF allows metadata to be attached to any stream in the document, such as information about embedded illustrations and the whole document.</p>	<p>EFFECTIVE EDISCOVERY Electronic information is usually accompanied by metadata, which is not found in paper documents, and can play an important part as evidence. For example the date and time a document was written could be useful in a copyright case.</p>
<p>10 DOCUMENT USABILITY Interactive objects and file attachments PDF files may contain interactive elements such as links, bookmarks, table of contents, annotations, form fields, video and 3D objects. Interactive elements help with individual document navigation, grouping and indexing documents together.</p>	<p>KNOWLEDGE TRANSFER A well-built PDF allows readers to understand a document with relative ease and accurately follow the instructions they are given. PDF makes documents comprehensible to both humans and machines by introducing consistent terminology, logical sentence structures, sections and paragraphs, graphics.</p>

EVALUATING SOLUTIONS

As there are countless applications of PDF, so there are solutions that help businesses generate it. From manual one-by-one generation, to out-of-the box packaged solutions, to in-house development bundled into your existing applications, your IT team has a lot to consider.

For some businesses, a manual or a pre-built out-of-the-box PDF solution may be doing a sufficient job. In many cases though, PDFs need to be created in massive amounts, in an unattended mode. In addition, your requirements may be tied to business processes that are unique to your company or your existing company legacy systems such as websites, CRM, BPM, ERP and ECM.

In such cases, building in-house, or starting from scratch may be the way to go. For example:

- Your content is not available in advance, but is rather calculated based on user input or real-time database information.
- You have massive numbers of files need to be batch-created in unattended mode (for instance, inserting watermarks on a number of pages).
- Content needs to be customized or personalized on the fly.

Using a developer library

With an in-house app, it's up to you to provide ongoing maintenance, upgrades and troubleshooting, as well as any needed end-user support. In addition, you need to build it for performance, features, and standards compliance.

When evaluating a solution, ask the following questions:

- Does it give you full control over building custom scenarios?
- Is it easy to get started for developers?
- Does it have clean licensing?
- Is it well documented?
- Is it supported and actively developed?
- Is it compliant with ISO and industry standards?
- Is it supported for the long-term?
- Does it have an active user community?
- Does it have all features that you need?
- Is it extendable to features needed in the long-run?
- Does it contain quality, fast and tested code?

By considering the factors above, your organization will be on the way to implementing a cost-effective, future-proof solution that helps you mitigate risk and run a sustainable business.

TAKING ACTION

Unstructured data continues to grow exponentially across enterprise networks. Digitization is driving many organizations to re-think their document management policies related to regulatory compliance and retention (risk mitigation). Knowing the company's strategic goals, coupled with tactical knowledge of the most up-to-date capabilities of PDF, will lead to a future-proof solution for archiving, search, usability and compliance.

Therefore, the duty of a business, IT or a dedicated records professional will be in establishing a well-structured PDF that stays optimized for processing regardless of how the technology and platforms around it change over time. As the PDF format continues to evolve and its usage grows in importance and complexity, the demand for sophisticated tools for business will increase as well.

ABOUT ITEXT

iText, is one of the most comprehensive and actively developed PDF developer libraries on the market. It is already an integral component of many in-house or OEM-led solutions at the largest companies. iText's library contains an endless number of features that can be integrated programmatically into any web, mobile or a cloud platform, such as creating compliant PDFs from data; splitting and merging; stamping content; filling and flattening forms and inserting digital signatures. As an active member of the PDF ISO committee, iText also differentiates itself as a thought leader and forward-looking vendor with customer value in mind.

To find out more about iText, visit: <http://www.itextpdf.com>