

## Grazie Mille, PDF! Das Format für alle Fälle.

#Datenaustauschformat, #ISO-Standards, #Daten- und Transaktionsdruck,  
#Barrierefreiheit, #Archivierung



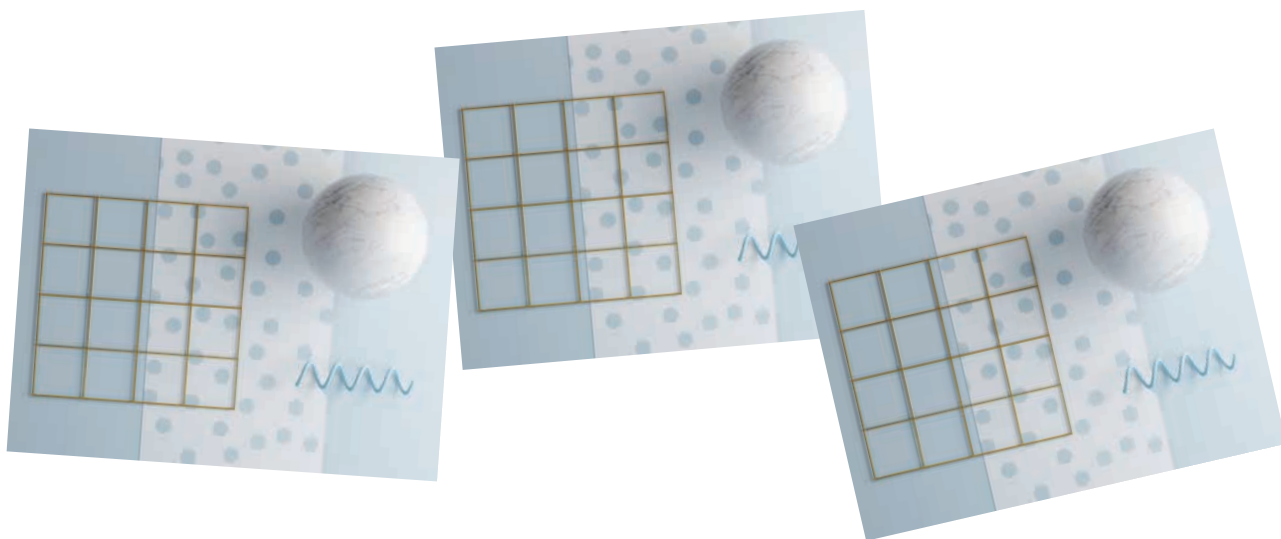
**Dietrich von Seggern**, Geschäftsführer der **callas software GmbH**, ist seit 1991 im Bereich der Druckvorstufe tätig. Der Diplom-Ingenieur ist Experte im Bereich Publishing und PDF. callas software entwickelt PDF-Technologien für das Publishing, Prepress, den Dokumentenaustausch und die Dokumentenarchivierung. Das Unternehmen ist Gründungsmitglied der PDF Association und engagiert sich von Anfang an im Vorstand des internationalen Verbandes.

[www.callassoftware.com/de](http://www.callassoftware.com/de)

PDF zählt weltweit zu den stark verbreiteten Formaten. Zahlreiche Unternehmen tauschen darüber zwischen Geschäftspartnern oder im eigenen Haus Informationen aus. Denn PDF bietet auf der einen Seite zahlreiche Möglichkeiten – auf der anderen ist das Format komplex. So ist es sinnvoll und notwendig, über verbindliche Qualitätsaussagen die Interoperabilität innerhalb und außerhalb eines Unternehmens sicherzustellen. Diese Qualitätsanforderungen werden über die PDF Standard-Formate definiert und ihre Einhaltung vereinfacht. Die Standards entlasten zudem von technischen Details und helfen bei der Klärung von Zuständigkeiten innerhalb der Prozesskette.

Um angesichts der vielen PDF-Erzeugungsprogramme und der zentralen Bedeutung von PDF eine zuverlässige PDF-Verarbeitung zu garantieren, wurden die Standards zusätzlich durch die ISO definiert. Eine oft gestellte Frage in diesem Zusammenhang lautet, warum nicht ein einziger Standard für „qualitativ gutes“ PDF ausreicht. Der Grund liegt in den verschiedenen Anwendungssituationen des Formats und das Ziel der ISO-Definitionen besteht darin, Anwendern ein für ihre Zwecke funktionales PDF zu ermöglichen. Ein Standard, der z. B. die Anforderungen für den Druck, die Vorgaben für die Archivierung oder die Barrierefreiheit kombinieren würde, wäre zwar ohne weiteres möglich, würde aber enormen Aufwand für die Erstellung valider PDFs nach sich ziehen. Eine Differenzierung erfolgt also hier mit Blick auf die unterschiedlichen Anforderungen und deren Umsetzung.

Der 2017 erschienene, jüngste ISO-Standard ist PDF selbst, das nun in Version 2.0 vorliegt. Die Definition umfasst fast 1.000 Seiten und enthält viele Detailverbesserungen und einige bedeutende Klarstellungen. Eine Reihe von Kapiteln wurde komplett neu verfasst, um die Verständlichkeit und Eindeutigkeit der Beschreibungen für die im Großen und Ganzen aber unverändert gebliebenen PDF-Konstruktionen zu verbessern.



## PDF/X – Standard für die Druckindustrie

Im Druck- und Prepress-Bereich sind PDF-Dateien das „Rohmaterial“. Dabei gibt es zahlreiche Anforderungen, die über das fehlerfreie Darstellen auf dem Monitor hinausgehen. Deshalb formulierte die Branche bereits kurz nach Erscheinen von PDF entsprechende Anforderungen und entwickelte PDF/X (X steht für eXchange) unter dem Dach der ISO. Dieser Standard wurde bereits 1999 veröffentlicht. Dieses Format ist mittlerweile in der Druckindustrie weitestgehend etabliert und wird von zahlreichen Softwarelösungen unterstützt. Auf PDF/X aufbauende Spezifikationen, z. B. von der Ghent Workgroup oder von der Schweizer Initiative PDFX-ready, definieren weitergehende Anforderungen für verschiedene Druckverfahren und -produkte.

Der technischen Weiterentwicklung wurde durch PDF/X-4 Rechnung getragen, das dem heutigen Stand der Technik entspricht. Der ebenfalls bereits verfügbare PDF/X-5-Standard deckt eher Spezialanforderungen ab. Derzeit arbeitet die ISO an PDF/X-6, in das Änderungen durch PDF 2.0 einfließen. Hierzu gehören seitenbasierte Output Intents, Black Point Compensation, Halftone Origin sowie spektrale Messdaten für Sonderfarben (CxF). Letzteres ermöglicht eine bessere Wiedergabe von Sonderfarben, beispielsweise im Unternehmenslogo, sowie eine konsistente Ausgabe auf unterschiedlichen Kanälen, wie Digitaldruck oder Druck auf unterschiedlichen Medien.

## PDF/A für die Langzeitarchivierung

Neben PDF/X spielt PDF/A in der Praxis eine wichtige Rolle. Der Standard wurde auf Betreiben der produzierenden Industrie

entwickelt, die ein anerkanntes, robustes PDF für die Archivierung ihrer Produktionsdokumente benötigte sowie von Archivierungseinrichtungen. Veröffentlicht wurde er Ende 2005 als internationaler Standard ISO 19005-1. Unternehmen und öffentliche Einrichtungen profitieren von PDF/A, da digitale Dokumente dauerhaft archiviert werden können. Während das Format zunächst in erster Linie als Ersatz für gescanntes Papier bzw. TIFFs in Archiven verwendet wurde, kommt es inzwischen überwiegend auch für digital erzeugte Dokumente zum Einsatz. Das Format ist vor allem in Europa stark verbreitet.

Mittlerweile existieren drei Normteile von PDF/A und unterschiedliche Konformitätslevel:

- Stufe B (Basic) garantiert eine eindeutige visuelle Reproduzierbarkeit der Inhalte. Sie lässt sich einfacher generieren als die übrigen Stufen, gewährleistet aber nicht zu 100 Prozent Textextraktion oder -durchsuchbarkeit. Eine problemlose Wiederverwendung des Inhalts ist nicht unbedingt gegeben. Gescannte Papiervorlagen lassen sich normalerweise ohne Weiteres in PDF/A-Dokumente dieser Konformitätsstufe konvertieren.
- Stufe U (Unicode) wurde mit PDF/A-2 eingeführt. Es handelt sich um eine Erweiterung der Konformitätsstufe B in der Form, dass der gesamte Text im Schriftzeichenstandard Unicode abgebildet ist, was eine Volltextsuche ermöglicht.
- Stufe A (Accessible/Zugänglich) umfasst sämtliche Anforderungen und verlangt zusätzlich Angaben zur inhaltlichen Struktur und zur korrekten Lesereihenfolge des Dokumentinhalts. Textinhalte müssen extrahierbar sein und die Struktur muss die natürliche Leseabfolge abbilden. Diese PDF/A-Stufe lässt sich ohne aufwändige Nachbearbeitung in der Regel nur bei der Konvertierung aus digitalen Ursprungsdokumenten erreichen.

Als dritter Normteil erschien PDF/A-3 im Jahr 2012: Er erlaubt das Einbetten beliebiger Dateien. So können beispielsweise digitale Akten komfortabel in einer Datei abgebildet werden. Ein anderes Beispiel sind ZUGFeRD-Rechnungen. In diesen sind der maschinell verarbeitbare XML-Datensatz mit den Rechnungsinformationen in einer vom Menschen lesbaren PDF/A-3-Datei integriert. Auch bei dieser Version wird derzeit an einer Aktualisierung des auf PDF 2.0 basierenden Standardteils gearbeitet, der als PDF/A-4 erscheinen würde. Ein wichtiger Teil dieses neuen Standards wird PDF/E sein.

### PDF/E fürs Ingenieurwesen

PDF/E wurde für das Ingenieurwesen entwickelt und 2008 als Standard veröffentlicht. Darin ist die Fähigkeit zur interaktiven 3-D-Darstellung integriert, sodass es besonders als Austauschformat für Konstruktionszeichnungen geeignet ist. In der Praxis ist dieser Standard jedoch kaum relevant. Die Fertigungsindustrie formuliert aber nach wie vor Bedarf an einem archivierbaren Basisformat, das – anders als PDF/A – auch 3-D-Modelle ermöglicht. So wird aktuell auch daran gearbeitet, das bisherige PDF/E-Format zu modernisieren und als Konformitätsstufe in PDF/A-4 zu integrieren.

### PDF/VT & PDF/VCR für variablen Daten- und Transaktionsdruck

PDF/VT (ISO 16612) ist das Austauschformat für variablen Daten- und Transaktionsdruck und wurde 2010 veröffentlicht. Es basiert auf PDF/X, bildet eine Alternative zu den Formaten PCL, PPML, AFP, etc. und adressiert mit der Individualisierung und dem Digitaldruck neue Trends in der Drucktechnologie.

Noch relativ jung ist PDF/VCR (ISO 16613) für variablen Daten- und Transaktionsdruck in Echtzeit. Getrieben aus der Praxis und im September 2017 veröffentlicht, erlaubt er das Definieren variabler Daten mit PDF-Templates basierend auf PDF/X-4. Er kommt bei Spezialanwendungen, die eine sichere Echtzeitverarbeitung benötigen, zum Einsatz. Ein Beispiel ist der Druck von Begleitschreiben für Kreditkarten, bei denen zur Ermittlung der Daten die Kreditkarte ausgelesen wird.

### PDF/UA für barrierefreie Dokumente

Im ISO-Standard PDF/UA steht „UA“ für universal accessibility (Barrierefreiheit). Die Anforderungen für eine PDF/UA-konforme Datei definieren, wie Texte, Bilder, Formulare und Ähnliches beschaffen sein müssen, damit Menschen mit Behinderungen – und Maschinen – sie nutzen können. Somit hilft PDF/UA, gesetzliche Vorgaben nach einem uneingeschränkten Zugang zu elektronischen Informationsangeboten – zum Beispiel in öffentlichen Einrichtungen, Versicherungen und Banken – zu erfüllen.

Eine technische Basis von PDF/UA sind die Anforderungen an Texte aus PDF/A-2u. Darüber hinaus enthält der Standard Anforderungen an die normalerweise nicht angezeigte Struktur eines Dokuments, also Lesereihenfolge und Gliederung in Überschriften, Absätze, Spalten, Tabellen und Alternativtexte für Bilder. Eine automatische, nachträgliche Generierung der Struktur ist ausgesprochen aufwendig, die Erzeugung von PDF/UA beginnt daher in der Regel bereits am Ursprungsdokument.

PDF/UA kann auch als Nachfolger der Konformitätsstufe „A“ aus PDF/A gesehen werden und stellt von allen hier vorgestellten PDF-Standards die höchsten Anforderungen an die Erzeugung. Während PDF/A-Dateien in Konformitätslevel „B“ aus fast allen PDF-Dateien erzeugt werden können, ist dies bei PDF/X nicht immer und bei PDF/UA nur sehr selten möglich.

### Ausblick: Werden PDF-Standards zukünftig kombiniert?

Um die Erzeugung standardkonformer PDFs so einfach wie möglich zu halten, werden die genannten PDF-Standards voraussichtlich weiterhin getrennt voneinander bestehen bleiben. In den ISO-Komitees wird aber daran gearbeitet, dass die nächsten Versionen, wo es immer möglich ist, exakt die gleichen Formulierungen verwenden, damit Software-Anbieter ohne großen Aufwand mehr als einen Standard unterstützen können. Die Standards sind aber bereits heute so verfasst, dass solche „doppelt validen“ PDF-Dateien ohne Weiteres möglich sind. Auch wenn einiges an Spezialwissen erforderlich ist, um mit PDF Erzeugungsprogrammen, wie Microsoft Word, Adobe InDesign und anderen, PDFs zu erzeugen, die mit mehreren Standards gleichzeitig konform sind, gibt es durchaus Produkte, die die nachträgliche Konvertierung – auch in automatisierten Abläufen – ermöglichen. ■