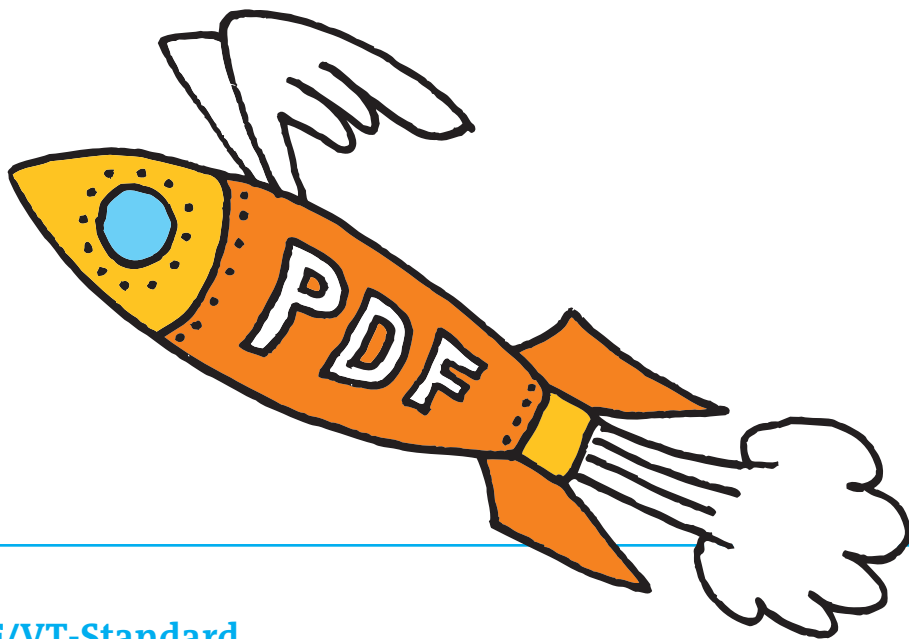


Technische Einführung in PDF/VT



 PDFlib® Whitepaper

Der PDF/VT-Standard

Der PDF/VT-Standard wurde 2010 als ISO 16612-2:2010 veröffentlicht. Der Standard ist »dazu entworfen, *Variable Document Printing* (VDP) in einer Vielzahl von Umgebungen zu ermöglichen«. PDF/VT-Dokumente enthalten die finalen Inhalte und die zugehörigen Metadaten, aber keine Variablen oder Templates. Der Vorgängerstandard ISO 16612-1:2005 spezifizierte das PPML/VDX-Format, das auf PDF 1.4 basiert. Diesem älteren Standard fehlten jedoch einige PDF-Funktionen und er beruhte auf Konstrukten außerhalb von PDF. Der moderne Standard PDF/VT basiert auf PDF/X-4 und PDF/X-5 und unterstützt den Funktionsumfang von PDF 1.6 einschließlich Transparenz, Ebenen und ICC-basiertem Farbmanagement. Zusätzlich zu den Anforderungen von PDF/X steuert der PDF/VT-Standard weitere Funktionen bei, um den Anforderungen an personalisierten Massendruck zu entsprechen. PDF/VT erlaubt schnelles Rastern (Rendering) von digitalen Druckdateien, indem es das PDF-Format um effizientes Ressourcen-Management ergänzt.

Da PDF/X Farbmanagement und Unterstützung für ICC-Profile enthält, hat auch PDF/VT diese Funktionen übernommen. Für Organisationen, die sich bisher auf schwarzweiße Drucksachen beschränkt haben, kann Farbmanagement einen hohen Einarbeitungsaufwand bedeuten, eröffnet jedoch auch eine Vielzahl von Möglichkeiten für neue farbige Druckprodukte und konsistente Farbproduktion.

PDF/VT zielt auf die beiden wichtigen Bereiche »*Variable Document Printing*« und »Transaktionsdruck« ab, die im Folgenden genauer beschrieben werden.

Variable Document Printing

Variable Document Printing, auch als Personalisierung oder Eins-zu-eins-Marketing bezeichnet, generiert individualisierte Seiten, d.h. manche oder alle der gedruckten Seiten sind unterschiedlich. Die variablen Inhalte kommen aus einer Datenbank und werden genutzt, um jedes Dokument anzupassen. VDP wird vorwiegend für das Marketing eingesetzt, da personalisierte Dokumente die Aufmerksamkeit des Empfängers schneller wecken. Hier einige Beispiele:

- ▶ Geschäftsbriefe mit variablen Namens- und Adressinformationen
- ▶ Personalisierte Broschüren mit variablen Text- und Grafikelementen, die nach einem Satz von Gestaltungsregeln ausgewählt werden. Daraus entstehen grafisch ansprechend gestaltete Unikate.
- ▶ Vollständig personalisierte Angebote, bei denen der eigentliche Inhalt anhand von Standort, Geschlecht, Alter und Kaufgewohnheiten oder anderen bekannten Daten über den Kunden variiert.

Transaktionsdruck

Transaktionsdruck erstellt zu Transaktionen wie Verkäufen, Kontobewegungen oder Versicherungsabschlüssen die zugehörigen Dokumente, also Rechnungen, Kontoauszüge oder Versicherungspolizen. Falls Werbe- oder Marketingbotschaften in diese Dokumente integriert werden, spricht man auch von Transaktional-promotional oder Transpromo. Gängige Beispiele für verschiedene Branchen:

- ▶ Finanzbranche: Ein Kontoauszug enthält maßgeschneiderte Angebote für den einzelnen Kunden.
- ▶ Versorgungswirtschaft: Eine Strom- oder Gasrechnung enthält individuelle Serviceangebote, angepasst an das Verbrauchsmuster des Kunden.
- ▶ Telekommunikationsbranche: Telefonrechnungen enthalten individuelle Rabattangebote anhand des Verbindungstyps und der Telefonnummern der im letzten Monat vom Kunden getätigten Telefonate.

Konformitätsstufen: PDF/VT-1, PDF/VT-2 und PDF/VT-2s

ISO 16612-2 spezifiziert drei Stufen der Konformität mit PDF-VT, die alle auf PDF 1.6 basieren:

- ▶ PDF/VT-1 wurde für den Austausch abgeschlossener Einzeldateien entwickelt und beruht auf PDF/X-4. Eine PDF/VT-1-Datei enthält alle erforderlichen Ressourcen für die Wiedergabe des Dokuments. Aus diesem Grund ist PDF/X-4 mit externem Output Intent nicht erlaubt.
- ▶ PDF/VT-2 ist für den Austausch mehrerer Dateien konzipiert und basiert auf einem der Standards PDF/X-4p, PDF/X-5g oder PDF/X-5pg. PDF/VT-2-Dateien können also auf externe Druckausgabebedingungen, externe Seiteninhalte oder beides verweisen. Ein PDF/VT-Dokument und alle referenzierten PDF-Daten und externen Druckausgabebedingungen werden zusammen als PDF/VT-2-Fileset bezeichnet.
- ▶ PDF/VT-2s ermöglicht die Ausgabe von Streams, damit sich die Generierung des Dokuments und das Rendering überlappen können. Ein PDF/VT-2s-Stream ist ein MIME-Paket, das ein oder mehrere PDF/VT-1-Dateien oder PDF/VT-2-Filesets (oder beides) enthält sowie alle referenzierten Dateien.

Technische Konzepte in PDF/VT

Dieser Abschnitt bietet einen Überblick über die technischen Konzepte, auf denen PDF/VT beruht.

Document Part Hierarchy

Die *Document Part Hierarchy* (DPart) beschreibt die Abfolge und Beziehung von Dokumenten oder Teilen von Dokumenten in einer PDF/VT-Datei. Üblich ist, dass die PDF/VT-Datei aus Teildokumenten für viele Empfänger besteht und jeder Dokumententeil die Seiten für einen einzelnen Empfänger enthält. Die DPart-Hierarchie kann aber nicht nur Seiten den jeweiligen Empfängern zuordnen, sondern auch komplexere Strukturen abbilden. Beispielsweise könnten die Empfänger nach der Postleitzahl in ihrer Adresse, die Postleitzahl wiederum nach Bundesland und die Bundesländer nach Staaten sortiert sein. Diese Art der Organisation eines Dokuments schafft eine Baumstruktur, die alle Seiten des Dokuments enthält. Die Elemente dieses Baums heißen DPart-Knoten, wobei jeder interne Knoten weitere DPart-Knoten enthält und jeder Blattknoten eine oder mehrere Seiten für einen Empfänger beschreibt.

Alternativ zum Zugriff über Seitennummern oder Seiten-Labels kann diese Hierarchie von Dokumentteilen in einer PDF/VT-Datei benutzt werden, um auf Seiten zuzugreifen. Die DPart-Hierarchie ist in einer PDF/VT-Datei zwingend erforderlich. Der optionale Eintrag *RecordLevel* in einem PDF/VT-Dokument beschreibt die Ebene in der DPart-Hierarchie, auf der sich die Datensätze für individuellen Empfänger befinden. Dies ist bedeutsam für die *Scope Hints* (siehe unten).

Document Part Metadata

Jeder Knoten in der Hierarchie der Dokumentteile von der Wurzel bis zu den Blättern im Dokumentbaum kann *Document Part Metadata* (DPM) enthalten. Solche Metadaten können Informationen über das Dokument eines einzelnen Empfängers und dessen Teile vermitteln. Insbesondere produktionsrelevante Informationen (z.B. Anzahl der Kopien für einen Dokumententeil) oder Informationen über den Empfänger (z.B. die zugehörige Postleitzahl), können in DPM abgelegt werden.

Der Standard PDF/VT beschreibt die generellen Methoden zur Unterbringung der *Document Part Metadata*, schreibt jedoch weder Schema noch Kodierung vor. Allerdings wurde der Standard im Hinblick auf das Metadaten-Schema im Job Definition Format (JDF) entwickelt, ein von der *International Cooperation for the Integration of Processes in Prepress, Press, and Postpress Organization* (CIP4) standardisiertes Jobticket-Format. Das für PDF/VT empfohlene Metadaten-Format heißt »ICS – *Common Metadata for Document Production Workflows*«. Weitere Informationen dazu finden Sie unter www.cip4.org.

JDF (oder andere) Produktions-Metadaten sind in PDF/VT nicht zwingend erforderlich, erweisen sich aber in JDF-fähigen Workflows als äußerst nützlich. Zusätzlich beschreibt der PDF/VT-Standard auch eine Methode zur Darstellung von DPM mit XML.

Optimierungen für wiederkehrende grafische Elemente

Druckelemente werden oft auf vielen Seiten mehrfach genutzt, z.B. ein Firmenlogo oder ein Produktbild, das auf allen Seiten des Mailings auftaucht. Die optimierte Verarbeitung dieser wiederkehrenden grafischen Inhalte stellt eine wichtige Strategie dar, um sowohl Dateigröße als auch Verarbeitungsgeschwindigkeit der Druckdatei zu optimieren. PDF unterstützt seit jeher das Konzept der XObjects, die die Dateigröße minimieren, indem sie die erforderlichen Daten eines Druckelements nur einmal in der Datei vorhalten. Diese Daten können dann von beliebig vielen Seiten (oder mehrfach auf einer Seite) referenziert werden. XObjects können Rasterbilder, Text und Vektorgrafik enthalten. Während XObjects in PDF dabei helfen, die Dateigröße zu optimieren, enthielt PDF bisher keine Vorkehrungen, um die Verarbeitung von sich wiederholenden Seiteninhalten zu beschleunigen. Ein PDF-Dokument enthält keine Informationen darüber, dass zum Beispiel ein Bild auf einer bestimmten Seite auf einer weiteren Seite im gleichen oder einem anderen Dokument wiederholt wird. PDF/VT erweitert das existierende Konzept von XObjects in PDF und fügt die folgenden Methoden zur Optimierung der Druckgeschwindigkeit ein:

- ▶ **Eindeutige Identifizierung:** XObjects können eine Identifizierung enthalten (genannt *GTS_XID*), die über alle Dokumente eindeutig ist. Diese Identifizierung kann von Cache-Implementierungen genutzt werden, die gleiche XObjects erkennen müssen. Einfach gesagt, die bereits in Job 1 genutzte Grafik, welche in Job 2 wieder auftaucht, muss nicht noch einmal gerastert werden, sondern die vorhandenen Rasterdaten können aus dem Cache übernommen werden.
- ▶ **Scope Hints und Umgebungskontext:** XObjects können Informationen darüber enthalten (genannt *GTS_Scope*), die beschreiben, welche Seiten oder Dokumente die Inhalte erneut nutzen. Auf diese Weise tragen die XObjects Informationen über die sinnvolle Lebensdauer ihrer Rasterdarstellung im Cache: Wird der Inhalt nur für den aktuellen Empfänger gebraucht, an anderer Stelle in der gleichen Datei oder im gleichen Datenstrom wiederverwendet, oder vielleicht überhaupt nicht mehr benötigt? Wird der Umgebungskontext (genannt *GTS_Env*) angegeben, kann das XObject auch für die globale Nutzung vorgesehen sein, das heißt, es wird in mehr als einer PDF/VT-Instanz genutzt. Es gibt keine Einschränkungen, was der *String* zum Umgebungskontext enthalten kann. So kann zum Beispiel ein Kunden- oder Jobname zur Beschreibung des Kontexts angegeben werden.
- ▶ **Encapsulation Hints:** Caching-Algorithmen müssen das Zusammenspiel eines XObject mit dem aufrufenden Kontext und vorhandenen Druckelementen auf der selben Seite (oder anderen Seiten, wenn mehrere Seiten auf demselben Bogen ausgegeben werden) einbeziehen. Wenn in einem XObject zum Beispiel die Farbe oder Linienstärke nicht festgelegt wird, sondern diese Eigenschaften anhand der in der Umgebung vorherrschenden Farb- und Linieneinstellungen variieren, dann ist das Zwischenspeichern der gerenderten Ergebnisse aufgrund des unterschiedlichen Erscheinungsbilds nutzlos. Eine ähnliche Situation tritt ein, wenn das XObject transparente Elemente enthält, so dass der vorhandene Hintergrund mit dem XObject zusammen berücksichtigt werden muss. Um das *Caching* von XObjects zu vereinfachen, führt PDF/VT das Konzept des *Encapsulated XObject* ein, das als solches (mit Hilfe des Eintrags *GTS_Encapsulated*) markiert werden kann. *Encapsulated XObjects* müssen bestimmte Regeln erfüllen, die das *Caching* erleichtern.

All diese Einträge sind optional: PDF/VT verlangt keine dieser Optimierungen für sich wiederholende grafische Inhalte. Wenn sie allerdings genutzt werden, ergeben sich erhebliche Geschwindigkeitsvorteile mit einem PDF/VT-fähigen RIP.

Bildschirmdarstellung und Druck von PDF/VT-Dokumenten

PDF/VT-Dokumente sind zunächst ganz normale PDF-Dateien mit bestimmten zusätzlichen Eigenschaften. Allerdings beeinflussen diese Erweiterungen für PDF/VT nicht die Darstellung: Die Seiten eines PDF/VT-Dokuments können mit einem normalen PDF-Viewer betrachtet und mit einem PDF-fähigen RIP – der nicht unbedingt PDF/VT unterstützen muss – ausgegeben werden, also beispielsweise mit Acrobat oder einem Viewer oder RIP eines anderen Herstellers. Trotzdem müssen manche Aspekte mit Vorsicht behandelt werden:

- ▶ Die Farbmanagementfunktionen in PDF/VT erfordern einen PDF-Renderer, der das Farbmanagement mit ICC-Profilen vollständig umsetzt, genauso wie das auch die Standards PDF/X und PDF/A erfordern. Moderne PDF-Viewer erfüllen diese Anforderungen an die Farbdarstellung von PDF/VT.
- ▶ Die *Document Part Hierarchy*, *Document Part Metadata* (DPM) und die Optimierungen für wiederkehrende grafische Inhalte haben keine Auswirkung auf die Darstellung in einem normalen PDF-Viewer.
- ▶ PDF/VT-2 erlaubt externe, referenzierte Seiteninhalte. In Acrobat muss die Unterstützung externer Referenzen sorgfältig konfiguriert werden, um den gewünschten Effekt zu erzielen.
- ▶ MIME-Pakete gemäß PDF/VT-2s sind keine PDF-Dokumente. Der Inhalt eines solchen Pakets muss erst ausgepackt werden, bevor das Dokument in einem normalen PDF-Viewer betrachtet werden kann.

Missverständnisse zu PDF/VT

Da in einigen Publikationen Fehlinformationen über PDF/VT verbreitet wurden, wollen wir die folgenden Punkte klarstellen:

- ▶ Trotz des Namens enthält PDF/VT keine Variablen: Alle Seiten enthalten oder referenzieren den endgültigen Inhalt, ohne irgendwelche Möglichkeiten für den Einsatz von Variablen, dynamischen Elementen oder Formatierung von Inhalten. Ebenfalls nicht korrekt ist die Aussage, dass Acrobat PDF/VT-Dokumente ohne die variablen Daten darstellt.
- ▶ Obwohl PDF/VT-2s für das Streaming bei der PDF-Erstellung ausgelegt ist, kann die Erstellung eines PDF/VT-Dokuments nicht gleichzeitig mit der Rasterung für den Druck erfolgen, denn PDF-Dokumente (unabhängig vom PDF/VT-Status) müssen komplett vorliegen, bevor das Rendering starten kann. Die einzige Ausnahme zu dieser Regel ist das für die Webausgabe linearisierte PDF. Aber diese Technik ist für die schnelle Darstellung der ersten Seite im Browser ausgelegt, nicht um Tausende

PDF/VT und PDF/A für die Archivierung

von Seiten zu drucken, während der Rest der Datei noch erstellt wird.

Das folgende Szenario ist jedoch möglich: Ein PDF/VT-Paket enthält mehrere Dokumente, die über das Netzwerk in einem MIME-Paket übertragen werden. Sobald alle erforderlichen Teile für ein Dokument innerhalb des Pakets verfügbar sind, kann das Rendern des Dokuments beginnen, während andere Dokumente des gleichen Pakets noch über das Netzwerk übertragen werden.

PDF/VT-1-Dokumente können gleichzeitig konform zu PDF/A für die Archivierung (ISO 19005) sein. Das heißt, dass ein Dokument für den Digitaldruck auch als PDF/A archiviert werden kann. PDF/VT-2 lässt sich nicht mit PDF/A kombinieren, weil PDF/A keine externen Referenzen erlaubt. Da PDF/A-1 auf PDF 1.4 basiert, müssen sich kombinierte PDF/VT-1- und PDF/A-1-Dokumente auf die Funktionen beschränken, die in PDF 1.4 vorhanden sind. Insbesondere Transparenzen und Ebenen lassen sich nicht nutzen. Diese Einschränkung gilt für PDF/A-2 und PDF/A-3 nicht mehr, da diese Standards auf ISO 32000-1 (also PDF 1.7) basieren.

PDFlib und PDF/VT

Die PDF/VT-Unterstützung in PDFlib, PDFlib+PDI und PDFlib Personalization Server (PPS) basiert auf der vorhandenen Implementierung von PDF/X-4 und PDF/X-5, die 2009 eingeführt wurde. Alle relevanten PDF/X-Funktionen und -Anforderungen in PDFlib betreffen auch PDF/VT. Das schließt Farbmanagement mit ICC-Profilen, *Output Intents* zur Charakterisierung der geplanten Druckausgabebedingungen und Regeln für den Umgang mit Farbe, Ebenen und Transparenz im Dokument mit ein. Da PDFlib bereits extern referenzierte Dokumentbestandteile für die Standards PDF/X-4p, PDF/X-5g und PDF/X-5pg unterstützte, stellte es eine solide Grundlage für die Implementierung von PDF/VT-2 dar. Darüber hinaus war auch die Basis für PDF/VT-2s bereits vorhanden, da PDFlib schon immer die Dokumentgenerierung im Arbeitsspeicher unterstützt und eines der wenigen Produkte ist, die die Erstellung und Ausgabe von PDF verschachteln können. PDFlib unterstützt auch die Ausgabe von Dateien, die gleichzeitig die Vorgaben von PDF/VT-1 und PDF/A erfüllen.

Bei der Entwicklung der PDF/VT-Funktionen in PDFlib standen einige Funktionen im Mittelpunkt, etwa die automatische Erstellung von *Unique Identifiers* und eine flexible Programmierschnittstelle zur Beschreibung der *Document Part Hierarchy* und der *Document Part Metadata (DPM)*. Da PDFlib nicht weiß, in welcher Beziehung mehrere einzeln generierte Dokumente zueinander stehen, kann der Anwendungsentwickler *Scope Hints* und *Encapsulation Hints* hinzufügen. Programmierbeispiele zur Erstellung von PDF/VT stehen im PDFlib Cookbook auf der PDFlib-Website zur Verfügung.

PDF/VT-2s kann auf Basis von PDF/VT-1 oder PDF/VT-2 generiert werden, indem die PDF/VT-Dokumente zusammen mit allen referenzierten (externen) Druckausgabebedingungen und PDF-Inhalten in einem MIME-Paket verpackt werden. Auch hierzu steht im PDFlib Cookbook ein Beispiel zur Verfügung.

Im PDFlib pCOS Cookbook zeigt ein Codebeispiel, wie sich DPM aus PDF/VT-Dokumenten mit der pCOS-Schnittstelle auslesen und entsprechend des Standards als XML darstellen lässt.



PDFlib GmbH

Franziska-Bilek-Weg 9
80339 München, Germany
Tel.: +49 • 89 • 452 33 84-0
support@pdflib.com
www.pdflib.com/knowledge-base/pdfvt

PDFlib GmbH ist auf die Entwicklung von PDF-Technologie spezialisiert. PDFlib-Produkte sind seit 1997 weltweit im Einsatz. Das Unternehmen berücksichtigt wichtige technologische Trends, etwa ISO-Standards für PDF. PDFlib GmbH vertreibt alle Produkte weltweit, wobei Nordamerika, Europa und Japan die wichtigsten Märkte darstellen.