

White Paper

Serverbasierte PDF-Erstellung: Grundlagen

Inhaltsverzeichnis

Das PDF-Format.....	2
Beschreibung.....	2
Vorteile des Dateiformats.....	2
PDF-Erstellung: Desktop- oder Serverlösung?	3
Vorteile von Serverlösungen	3
Anforderungen an PDF-Server	4
Alternative Wege der serverbasierten PDF-Erstellung.....	5
Verteilung der PDF-Dateien	7
Beispielanwendung: PDF Render Center	7

Das PDF-Format

Nahezu jeder Computer-Nutzer kennt heute das PDF-Format. E-Mail-Nachrichten mit einem Anhang in Form einer PDF-Datei verdrängen heute vielfach Papierdokumente. So ist es zum Beispiel üblich, dass man die Rechnung für eine Flugbuchung im Internet innerhalb von Sekunden per E-Mail als PDF-Dokument erhält oder dass Telefongesellschaften die monatliche Rechnung als PDF-Datei elektronisch versenden und für die Zusendung der „originalen“ Papierrechnung manchmal sogar eine zusätzliche Gebühr verlangen.

Beschreibung

Das Portable Document Format (PDF) von Adobe® ist ein Dateiformat, mit dem die Anzeige von Dokumenten unabhängig von der ursprünglichen Anwendungssoftware sowie von der Hardware und vom Betriebssystem, die für die Erstellung der Dokumente verwendet wurden, erfolgen kann. Zusätzlich zum statischen Design der Seiten kann ein PDF-Dokument auch interaktive Elemente enthalten, die nur in elektronischer Form möglich sind: PDF unterstützt verschiedene Typen sogenannter Annotations wie zum Beispiel Textkommentare, Hyperlinks, Änderungen oder Markierungen, Dateianlagen, Sounddaten und Filme. Ein PDF-Dokument kann (in bestimmten Grenzen) seine eigene Benutzeroberfläche bei der Anzeige definieren. Eingaben durch Tastatur oder Maus können je nach Einstellung Aktionen auslösen, die durch PDF-Objekte genau beschrieben werden. Das PDF-Dokument kann auch interaktive Formularfelder enthalten, die vom Anwender ausgefüllt werden können. Außerdem ist der Export von Feldinhalten oder der Import dieser Inhalte aus anderen Anwendungen möglich. Zusätzlich zu den Angaben zum Design enthalten PDF-Dokumente häufig auch Identifikationsmerkmale und Details zur logischen Struktur, die das Durchsuchen, das Bearbeiten oder die Entnahme von Inhalten zur weiteren Verwendung an anderer Stelle ermöglichen. Schließlich kann der Umfang der Verwendung der fertigen PDF-Datei durch Sicherheitseinstellungen gesteuert werden, die man bei der Erstellung der Datei definiert.

Vorteile des Dateiformats

Mit der Verbreitung von elektronischen Dokumenten und der damit einhergehenden Verdrängung traditioneller Papierdokumente wächst zum einen die Anzahl digitaler Dokumentenformate. Zum anderen erweitern sich aber auch deren Funktionen und Vorteile. Jede Person, Organisation oder Institution und jedes Unternehmen muss nach der Entscheidung für ein - vollständiges oder teilweises - digitales Dokumenten-Management zunächst eine Entscheidung über das Datenformat treffen, das künftig verwendet werden soll. Zu Recht wird in diesem Zusammenhang von vielen Fachleuten empfohlen, das PDF-Format in die Überlegungen einzubeziehen. Wo liegen dessen Vorteile?

Vor 10 Jahren sprachen zunächst die Offenheit des Formats und eine umfassend dokumentierte Spezifizierung und damit Standardisierung für PDF. Darüber hinaus wurden die Flexibilität der technischen Voraussetzungen und die möglichen Funktionalitäten als Argumente für PDF genannt. Dies alles ist auch heute noch richtig und wichtig, aber PDF ist darüber hinaus mittlerweile zu einem Standard gereift. Das Dateiformat ist sogar als weltweiter Standard etabliert, spätestens seit seiner offiziellen Ernennung zum ISO-Standard durch die *International Organization for Standardization* im Juli 2008.

Das PDF-Format war bereits ohne diesen „offiziellen Ritterschlag“ sehr populär und fest etabliert, was für seine fortschreitende Verbreitung von Bedeutung ist und diese Verbreitung weiter forciert. Man kann sich heutzutage kaum einen Computer vorstellen, auf dem keine PDF-Software installiert ist: Mindestens zum Lesen und Ausdrucken verwendet nahezu jeder Computer-Anwender das PDF-Format. Und dies unabhängig vom installierten Betriebssystem. Die Anzahl an Software-Produkten zum Erstellen, Bearbeiten und Anzeigen von PDF wird weiter wachsen. Dadurch ist sichergestellt, dass das wichtigste Problem von digitalen Dokumenten - die problemlose Übertragbarkeit - gelöst werden kann: Der Inhalt eines Dokuments kann auf jedem Computer wegen der Standardisierung des Formats mindestens gelesen werden, und dies auch eine lange Zeit nach seiner Erstellung.

Da Dateien in beliebigen Formaten in PDF-Dokumente eingebettet werden können, wird die Abbildung der unterschiedlichsten Abläufe in einem Dokumenten-Workflow erleichtert: Es ist viel einfacher, nur eine Datei für einen Vorgang aufzubewahren und zu verwalten als mehrere miteinander verbundene Dateien. Zudem können XML-Metadaten in einem PDF-Dokument verwendet und mit unterschiedlichen Objekten in dem Dokument verbunden werden. Dies erweitert die Möglichkeiten zur Integration von nicht visuellen Daten in ein PDF-Dokument. Großes Augenmerk wurde beim PDF-Format auch auf die Sicherheit der Dokumenteninhalte gelegt. Bei der Arbeit mit PDF-Dokumenten kann vom Verfasser des Dokuments und vom Softwareentwickler, der PDF in seine Anwendungen integriert, zum Beispiel der zu verwendende Verschlüsselungsalgorithmus festgelegt werden, mit dem die Dateien vor unberechtigten Zugriffen geschützt werden sollen. Das PDF-Format erlaubt zudem die Integration von digitalen Unterschriften (Signaturen) unmittelbar in die PDF-Datei.

PDF-Erstellung: Desktop- oder Serverlösung?

Wer sich aktuell auf dem Softwaremarkt nach einer guten Lösung für das Konvertieren von Dokumenten in PDF-Dateien umschaute, wird auf eine relativ große Anzahl an Produkten treffen. Dabei bilden die Produkte, mit denen die PDF-Erstellung nur auf dem einen Computer erfolgen kann, auf dem die PDF-Software selbst installiert ist (Desktop-Produkte) zahlenmäßig eindeutig die Mehrheit. Sie sind in der Regel preisgünstiger als Serverlösungen. Bei Letzteren übernimmt ein zentraler Rechner beziehungsweise eine zentrale Software, der Server, die Konvertierungsaufgabe für die dezentralen Clients. Serverbasierte Lösungen sind vor allem mit Blick auf die Erstellung, Verarbeitung und Verteilung von großen Mengen von PDF-Dateien für größere Benutzergruppen deutlich leistungsfähiger. Hinsichtlich der Qualität der erstellten PDF-Dateien (originalgetreue Darstellung der Ursprungsdatei) unterscheiden Serverlösungen sich in der Regel nicht von den Desktop-Versionen. Ihr Preis liegt wegen des speziellen Einsatzzwecks über dem für die Desktop-Pendants desselben Herstellers.

Vorteile von Serverlösungen

Serverlösungen haben wie erwähnt bei einer großen Zahl an regelmäßig umzuwandelnden Dokumenten und/oder bei einer großen Benutzergruppe einige Vorteile. Die wesentlichen Gründe dafür liegen zum einen in der üblicherweise einfacheren netzwerkweiten Einrichtung der Software. Zum anderen sprechen die Funktionen zur zentralen Administration der PDF-Erstellung und die **Möglichkeiten zur komfortablen Verteilung der erstellten Dateien im Netzwerk** für eine Serverlösung.

Bei den meisten serverbasierten PDF-Produkten ist die **Installation von Software auf jedem einzelnen Arbeitsplatzrechner nicht erforderlich**. Dann muss eine PDF-Software lediglich auf

dem Server-Rechner installiert werden und danach wird den Arbeitsstationen der Zugriff auf die Software per Freigabe im Netzwerk eingeräumt. Es gibt aber auch Lösungen, bei denen auf jedem Arbeitsplatzrechner PDF-Software oder mindestens ein PDF-Druckertreiber installiert werden muss.

Noch wichtiger als eine komfortable Installation ist die Möglichkeit von Serverlösungen, **die Parameter für die PDF-Erstellung zentral vorgeben und steuern** zu können. Auf diese Weise können die PDF-Einstellungen, zum Beispiel zu den Bearbeitungsrechten, zum Zugriffsschutz auf die erstellten Dateien oder zur Dateigröße, für die Benutzer vereinheitlicht werden. Häufig können vom Administrator auch Benutzergruppenprofile definiert werden, in denen die PDF-Parameter je nach dem konkreten Einsatzzweck der PDF-Datei (zum Beispiel Verwendung für die Bildschirmanzeige im Internet, für die Druckvorstufe, für den Konzeptdruck auf Papier, mit oder ohne Möglichkeit zur Bearbeitung/Kommentierung) zusammengestellt und abgespeichert werden. Der Benutzer muss beziehungsweise kann dann nur das passende Profil auswählen, ändert aber nicht die einzelnen PDF-Einstellungen. Manche Software-Hersteller gehen für die Bereitstellung von vordefinierten Zusammenstellungen von PDF-Parametern auch einen gänzlich anderen Weg: Entweder wird für jede Zusammenstellung ein eigener PDF-Drucker eingerichtet oder es werden unterschiedliche E-Mail-Adressen beziehungsweise Verzeichnisse für eingehende Konvertierungsaufträge angelegt. Wenn für jede Zusammenstellung von Parametern ein eigener Drucker eingerichtet wird, so ist dieser im allgemein bekannten Drucken-Dialog einer Anwendung im Zweifel schneller zu finden als ein Benutzerprofil in einem anderen Dialogfenster.

Die Anbieter von serverbasierter PDF-Software unterscheiden sich hier teilweise auch danach, ob der einzelne Benutzer an seinem Arbeitsplatzrechner die PDF-Einstellungen selbst ändern oder nur die Vorgaben und Profile einsetzen kann, die der Administrator festgelegt hat.

Falls es gewünscht ist, können außerdem dank der zentralen Konvertierung alle Dokumente **auf einheitlich gestalteten Dokumentenvorlagen und Layouts** (mit einem Unternehmenslogo, mit Kontaktdaten oder sonstigen Vorgaben) ausgegeben werden. Weitere Vorteile von Serverlösungen sind:

- Die **zentralisierte Update- und Lizenzverwaltung**
- Für den **Einsatz in** bestehenden oder neu einzuführenden **Dokumenten-Management-Systemen** (DMS) bieten leistungsfähige serverbasierte Lösungen in der Regel bereits in der Basisausführung nützliche Funktionen (zum Beispiel zur Verteilung von Dokumenten) und Schnittstellen (zum Beispiel Webdienste). Um die Zusammenarbeit mit DMS optimal steuern zu können, bieten flexible Anbieter von PDF-Servern zudem Möglichkeiten zur umfassenden Anpassung ihres Produkts an die individuellen Erfordernisse des Kunden.

Anforderungen an PDF-Server

Grundsätzlich müssen serverbasierte Lösungen dieselbe Anforderung erfüllen wie die Desktop-Lösungen, nämlich einen mit Blick auf die jeweilige Aufgabenstellung optimalen Mix aus den folgenden Eigenschaften: **wenige Umwandlungsfehler** und **hohe Geschwindigkeit** bei der Konvertierung, **umfangreiche Funktionen** zur Festlegung der PDF-Einstellungen, **geringe Dateigröße** und nicht zuletzt eine **gute Darstellungsqualität**, auch bei der Ausgabe auf Papier.

Aus den grundsätzlichen technischen Vorteilen von PDF-Serverlösungen ergeben sich zusätzlich aber noch die folgenden Anforderungen an ein gutes Produkt: Das Rollout sollte möglichst **keine Installation von Software auf jeder Arbeitsstation** erfordern und die PDF-Einstellungen sollten komfortabel und umfassend **zentral administrierbar** sein. Drittens ist eine **flexible und leistungsfähige Verteilung** der erstellten PDF-Dateien an die Adressaten (die Verfasser des Konvertierungsauftrags und sonstige involvierte Personen) wichtig.

Daneben ist auch **das angebotene Preis- und Lizenzmodell** von Bedeutung: Gut sind flexible, modular aufgebaute Modelle, bei denen vom Kunden im Kern nur die Funktionen bezahlt werden müssen, die tatsächlich benötigt werden. Zudem sollte der Preis auch von der Anzahl der Anwender der Lösung im Netzwerk und vom Datenverkehrsaufkommen (Anzahl an konvertierten Dokumenten in einem bestimmten Zeitraum) abhängen.

Gute Serverlösungen erlauben **Multithreading**, also die gleichzeitige Abarbeitung mehrerer Operationen innerhalb der Anwendung. Damit kann die Prozessorauslastung verbessert und die Leistung von Multiprozessor-Computern erhöht werden. Im Ergebnis läuft damit die Erstellung und Verteilung einer großen Zahl von PDF-Dateien schneller ab als bei einer Software, die nicht multithread-fähig ist.

Aus technischer Sicht muss eine Serverlösung darüber hinaus aus den folgenden **Komponenten** bestehen: ein Systemdienst, der die gleichzeitige Verarbeitung von Dokumenten steuert, die Komponente(n) zur Umwandlung der Dokumente in PDF sowie ein Administrationsprogramm zur Festlegung der Einstellungen und zur Steuerung der Konvertierung und Verteilung.

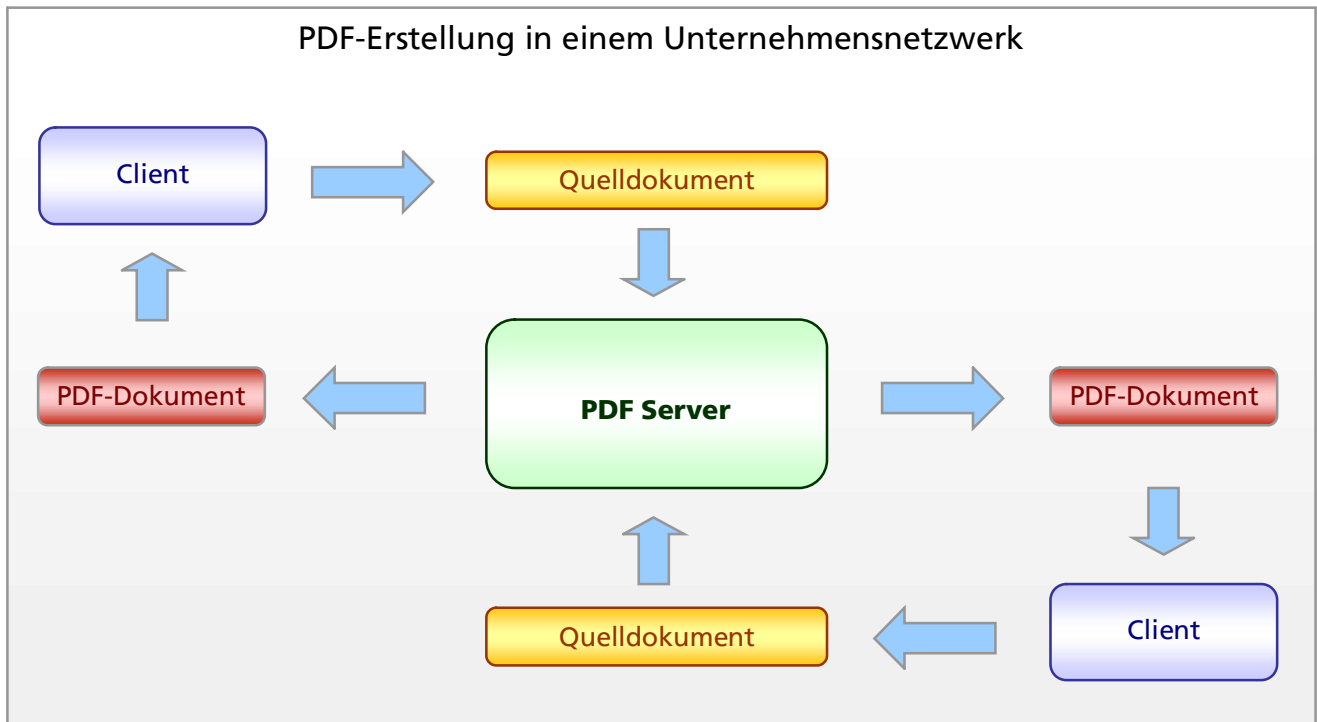
Abschließend sind die **Systemvoraussetzungen** für PDF-Serverlösungen zu erwähnen. Die Programme können in der Regel nicht nur auf „echten“ Server-Betriebssystemen wie Windows Server 2003 und 2008 eingesetzt werden, sondern auch unter Windows XP (Professional) oder Vista. Es gibt darüber hinaus auch Lösungen für Linux oder Apple Macintosh. Clientseitig webbasierte Lösungen existieren ebenfalls, sie können auf Windows-, Linux- und Mac-Rechnern eingesetzt werden.

Alternative Wege der serverbasierten PDF-Erstellung

Sowohl bei der Erteilung des Konvertierungsauftrags durch den Client als auch bei seiner Verarbeitung durch den Server gibt es unterschiedliche Vorgehensweisen:

Erteilung des Konvertierungsauftrags

Um bei einer Serverlösung eine PDF-Datei zu erstellen, geht der Benutzer in den meisten Fällen wie bei den verbreiteten desktopbasierten Produkten vor: Aus der Anwendung, in der eine PDF-Datei erstellt werden soll, wird am Arbeitsplatzrechner auf einem virtuellen Drucker „gedruckt“. Dieser PDF-Drucker wird bei der Installation der Serverlösung als zusätzlicher Drucker (Treiber) installiert und danach im Netzwerk freigegeben. Er gibt das Ergebnis des Druckvorgangs nicht auf Papier aus, sondern als PDF-Datei. Der Druck- beziehungsweise Konvertierungsauftrag kann bei Verwendung eines PDF-Druckertreibers für beliebige Anwendungen und Quellformate erteilt werden, sofern die Ursprungsanwendung über eine Druckfunktion verfügt.



Diesen Weg verwenden die meisten bekannten Serverlösungen. Es sind aber auch Produkte bekannt, bei denen die Konvertierungsaufträge als E-Mail an eine bestimmte Adresse gesandt werden, bei denen ein Upload der Ursprungsdatei über den Webbrowser erfolgt oder bei der die Originaldatei in einem freigegebenen Netzwerkordner abgelegt und automatisch in eine PDF-Datei umgewandelt wird. Auch Kombinationen dieser Alternativen kommen vor. Bei der PDF-Erstellung ohne Druckertreiber ist in der Regel die Zahl der konvertierbaren Dateiformate auf diejenigen begrenzt, die der jeweilige Hersteller in seiner Software zur Verfügung stellt.

Verarbeitung des Konvertierungsauftrags

Bei den Lösungen, die auf der Verwendung eines virtuellen Druckers basieren, kann der erste Konvertierungsschritt auf dem Arbeitsplatzrechner erfolgen. Aus der Ursprungsdatei oder aus dem Datenstrom wird hier eine (temporäre) Metadatei erzeugt und an den Server gesandt. Auf diesem erfolgt dann die Erstellung der PDF-Datei.

Von Bedeutung ist in diesem Zusammenhang, welchen Weg die Lösungen hier hinsichtlich der einzubettenden Schriften gehen: In Abhängigkeit davon, ob diese bereits am Arbeitsplatzrechner oder erst am Server eingebettet werden, müssen alle relevanten Schriften auch auf dem Server vorhanden sein oder auch nicht. Wenn die Einbettung am Arbeitsplatzrechner erfolgt, benötigt der Server die verwendeten Schriften nicht, bei der Einbettung am Server müssen dementsprechend dort dieselben Schriften wie auf den Arbeitsplatzrechnern vorhanden sein.

PDF kennt die folgenden 14 Standardschriften: Times-Roman, Helvetica, Courier, Symbol, Times-Bold, Helvetica-Bold, Courier-Bold, ZapfDingbats, Times-Italic, Helvetica-Oblique, Courier-Oblique, Times-BoldItalic, Helvetica-BoldOblique, Courier-BoldOblique. Sofern andere, tatsächlich

verwendete Schriften nicht eingebettet werden können, weil sie auf dem Arbeitsplatzrechner oder auf dem Server nicht installiert sind, werden die verwendeten Schriften durch ähnliche Standardschriften ersetzt, was in der Regel zu unerwünschten Ergebnissen führt - überlappende Buchstaben und/oder falsche Buchstabengröße.

Wird ein Dokument an dem Computer konvertiert, an dem auch die Ursprungsdatei erstellt wurde, erfolgt die PDF-Erstellung hinsichtlich der Schriften in der Regel ohne Schwierigkeiten, sofern alle Schriften dort lokal und korrekt installiert sind. Eine korrekte Einbettung der vorhandenen Schriften erfolgt deshalb für Desktop-Lösungen in der Regel problemlos und für Server-Lösungen immer dann, wenn sie wie bei den meisten PDF-Druckertreiber-Lösungen am Arbeitsplatzrechner erfolgt. Bei den Lösungen, bei denen die Ursprungsdatei per E-Mail an den Server versandt, per Upload geladen oder in ein Verzeichnis gelegt wird, müssen die Schriften immer auch auf dem Server installiert sein. Dies kann erhöhten Steuerungs- und Kontrollaufwand bedeuten, wenn an den Arbeitsplatzrechnern die Möglichkeit besteht, eigene Anwendungen und/oder Schriften zu installieren.

Verteilung der PDF-Dateien

Da bei Serverlösungen die Erstellung der PDF-Dateien aus Sicht der Clients „entfernt“ erfolgt, sind die von den Produkten angebotenen Wege zur Verteilung der fertigen Dateien von Bedeutung. Folgende Verteilungsoptionen sind möglich: Ablage in einem freigegebenen Netzwerkordner, in einem lokalen Verzeichnis auf dem Arbeitsplatzrechner, auch auf virtuellen Laufwerken, Versand per E-Mail, Webdownload oder über FTP-Server. Sollen die Dokumente auch auf Papier ausgegeben werden, ist dies über Netzwerkdrucker oder lokale Drucker möglich.

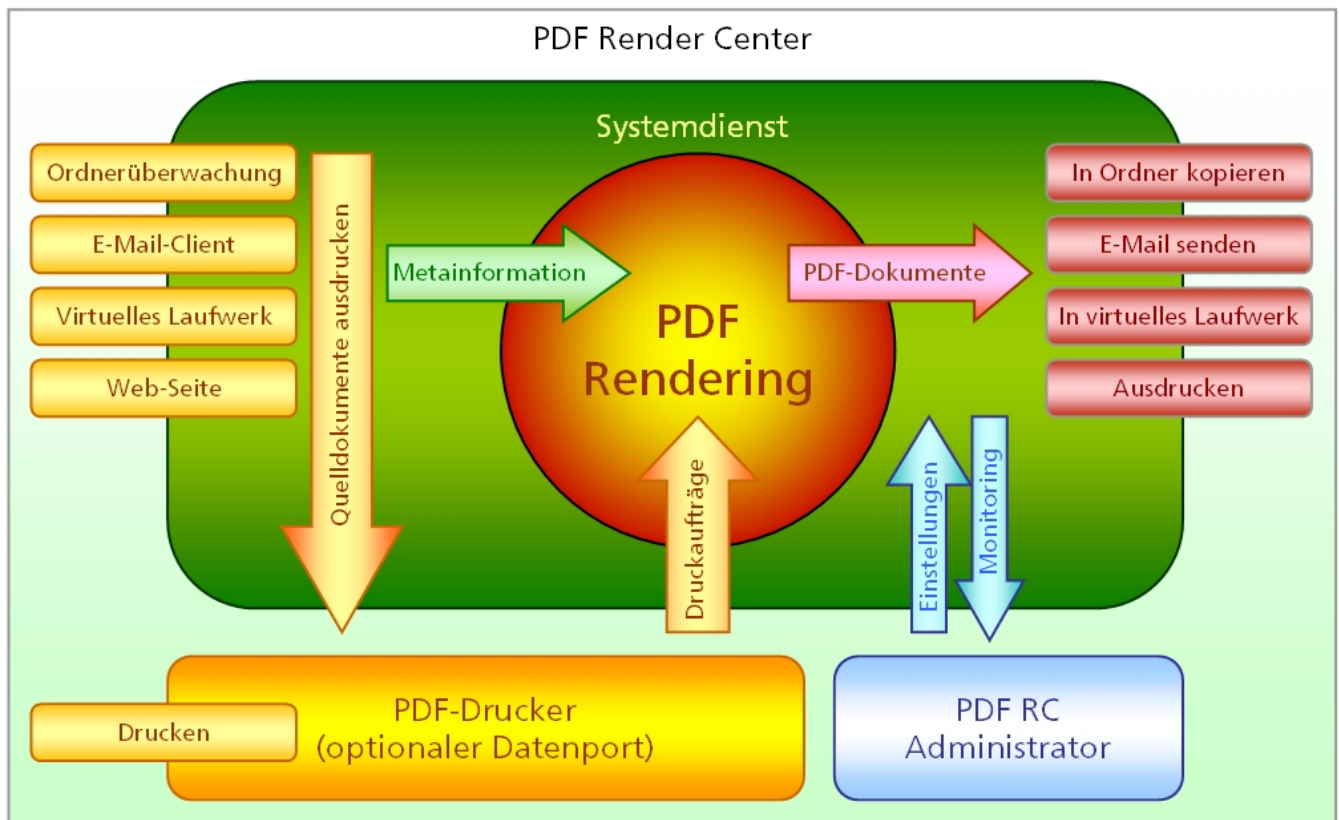
Beispielanwendung: PDF Render Center

Eine mögliche Serverlösung ist das [PDF Render Center](#) von soft Xpansion. Den Kern der Anwendung bildet ein Systemdienst von Windows, hinzu kommen ein virtueller Drucker (PDF-Drucker) und ein Administrationsprogramm für die Einstellungen. Die Lösung kann problemlos in bestehende Dokumenten-Management-Systeme integriert werden. Sie ist in Unternehmen aller Größen einsetzbar und ermöglicht die Vermeidung von Papierdokumenten sowie die automatische Verwaltung mehrerer Tausend Dokumente in einem einheitlichen Format. PDF Render Center ermöglicht wie alle serverbasierten Lösungen eine effizientere Verarbeitung einer großen Zahl an Dokumenten bei gleichzeitig reduzierten oder sogar entfallenden Kosten für die systematische Bearbeitung, für das Kopieren und für die Archivierung. Weitere Vorteile sind: geringere laufende Betriebskosten, Förderung der Kundenbeziehungen sowie die verbesserte Zusammenarbeit mit der eingesetzten Standard-Software, mit Toolkits und mit Individualsoftware, die in Content-Management-Produkte, ERP-Lösungen und sonstige vorhandene Systeme integriert wird.

Der Kern des PDF Render Center ist ein Systemdienst, der die Prozesse zur gleichzeitigen Bearbeitung von Dokumenten steuert. Der Dienst nutzt die Möglichkeiten von Multiprozessor-Systemen aus. Außer dem Systemdienst installiert PDF Render Center einen oder mehrere virtuelle Drucker (PDF-Drucker), die sowohl lokal als auch als Netzwerk-Drucker verwendet werden können. Die Administration des Render Center-Diensts erfolgt mit einer speziellen Applikation (PDF RC Administrator). Mit ihr steuert der Systemadministrator die Software und überwacht den Output.

Die wichtigsten Funktionen

PDF Render Center bietet die folgende Palette an Aktionen mit Dokumenten in den unterschiedlichsten Formaten:



- PDF- und PDF/A-Erstellung aus beliebigen druckfähigen Formaten: zum Beispiel Microsoft Office-Formate, formatierter (RTF) und unformatierter Text, HTML-Dateien und Webseiten, Bilder (TIFF, JPEG, PNG, GIF, BMP)
- Import von Microsoft Office-Metainformationen: Verweise, Inhaltsverzeichnis (TOC), Formularfelder und Steuerelemente, Kommentare, aus Word-, Excel-, PowerPoint-Dokumenten in PDF-Dokumente
- Konvertierung von PDF in PDF/A-1a und 1b
- Export von PDF-Dokumenten in ASCII-Text (ANSI oder Unicode) oder Bilder (TIFF, JPEG, PNG), inklusive mehrseitigen TIFFs, sowie auch CMYK

Grundsätzlich kann PDF Render Center beliebige Datenformate in PDF konvertieren. Die einzige Voraussetzung ist, dass auf dem Server, auf dem PDF Render Center installiert ist, auch die Software eingerichtet ist, die dieses Format auf dem PDF-Drucker ausdrucken kann. So ist zum Beispiel für die Konvertierung der Microsoft Office-Formate ein installiertes Microsoft Office auf dem Server erforderlich. Die anwendungsspezifischen Anforderungen sind zusätzlich:

- Import von Metainformationen aus Microsoft Office-Dokumenten wird ab Office 2000 unterstützt, empfohlen wird aber eine Version ab Office 2003
- Import von HTML-Dateien erfordert den Internet Explorer Version 6.0, empfohlen wird aber die Version 7.0 oder höher

Im Netzwerk sind mehrere Kommunikationsvarianten zwischen dem Arbeitsplatzrechner und PDF Render Center möglich, zum Beispiel Internet Explorer oder Outlook. Alle Varianten basieren auf einer Standardsoftware und erfordern keine zusätzliche Software und keine speziellen EDV-Kenntnisse.

