

Tabellen aus PDF-Dokumenten in anderen Programmen verwenden

Gibt es ein Programm, mit dem ich PDF-Dateien in eine Excel-Datei umwandeln kann?


Sehr häufig wird die Frage gestellt: Gibt es ein Programm, mit dem ich ein PDF-Dokument in eine Excel-Tabelle konvertieren kann oder Tabellen aus einem PDF in eine andere Anwendung kopieren kann? Diese Frage wird oft von Anwendern gestellt, die Texte aus vorhandenen PDF-Dokumenten in anderen Programmen weiterverarbeiten sollen, wenn z. B. die Originaldatei, aus der das PDF-Dokument erzeugt wurde, nicht mehr vorhanden ist, oder aber der Fragesteller will die Daten einfach in Excel bearbeiten und korrigieren können, ohne dass er den Text erneut abtippen muss.


Die Antwort auf die Frage lautet: Ja. Und das Beste dabei ist, dass die Anwender, die bereits mit ACROBAT XI PRO arbeiten, gar kein neues Programm kaufen müssen, denn sie haben ja schon das entsprechende Programm auf der Arbeitsstation installiert. Nämlich ACROBAT XI PRO. Schon seit vielen Versionen ist Acrobat in der Lage, ein PDF-Dokument in ein Excel-Dokument zu konvertieren. War das Ergebnis in früheren Versionen nicht gerade zufriedenstellend, so wurde es von Version zu Version immer wieder verbessert, bis es heute – mit der aktuellen Version XI – sehr gute und brauchbare Ergebnisse liefert.

Wie kann ich eine Tabelle aus einem PDF-Dokument in eine Excel-Arbeitsmappe oder in ein InDesign-Dokument kopieren?

Wenn Sie in Acrobat den Befehl zum Speichern einer PDF-Datei als MICROSOFT EXCEL-ARBEITSMAPPE aufrufen, werden immer alle Seiten des aktuell geöffneten PDF in die Zielformat gespeichert. Aber nicht immer muss der gesamte Inhalt eines PDF-Dokuments in eine Excel-Datei gespeichert werden. In den meisten Fällen handelt es sich nur um eine Tabelle, die zur weiteren Bearbeitung benötigt wird. Acrobat bietet dem Anwender die Möglichkeit, einzelne Bereiche in einem PDF zu markieren und zu konvertieren.

Ausgewählten Text in einem PDF-Dokument in Zwischenablage kopieren

Beim Kopieren von Texten oder Tabellen in einem PDF-Dokument kommt es darauf an, wie die Daten in Acrobat in den Arbeitsspeicher kopiert werden. Die meisten Acrobat-Anwender kennen und nutzen die Möglichkeit, Text in einem PDF-Dokument mit dem AUSWAHLWERKZEUG FÜR TEXT UND BILDER  in ACROBAT XI PRO zu markieren, ihn zu kopieren und über die Zwischenablage in ein anderes weiterverarbeitendes Programm einzufügen.

Allerdings kann das Aktivieren bestimmter Bereiche in Acrobat, wie in unserem Beispieldokument, schon zur Herausforderung werden. Wenn Sie beispielsweise versuchen, die Tabelle einfach zu markieren, indem Sie bei ausgewähltem AUSWAHLWERKZEUG FÜR TEXT UND BILDER  über den Text fahren, um

11 Versionsinfo

Die Beschreibungen wurden anhand der Version ACROBAT XI PRO erstellt.

Inhalte aus PDF-Dokument in anderen Programmen verarbeiten

Acrobat XI Pro kann PDF-Dokument als Excel-Arbeitsmappe speichern

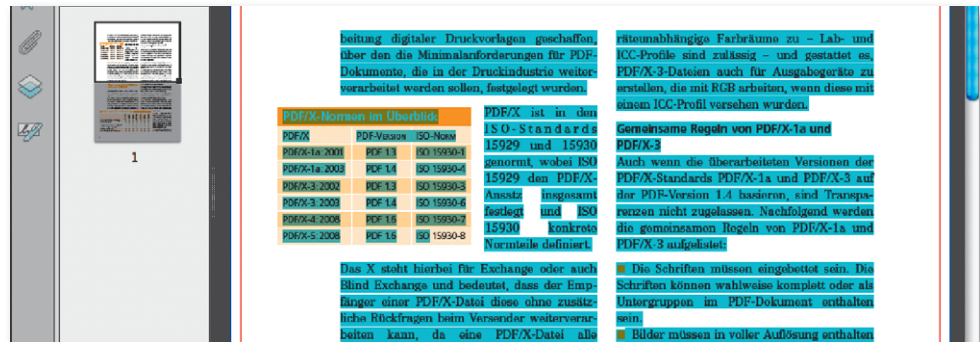
Alle Seiten einer PDF-Datei werden ins Zielformat gespeichert

Datenaustausch über Kopieren und Einfügen

Texte/Tabellen richtig markieren

ihn zu markieren, werden Sie schnell merken, dass gleichzeitig auch der umgebende Text mitmarkiert wird.

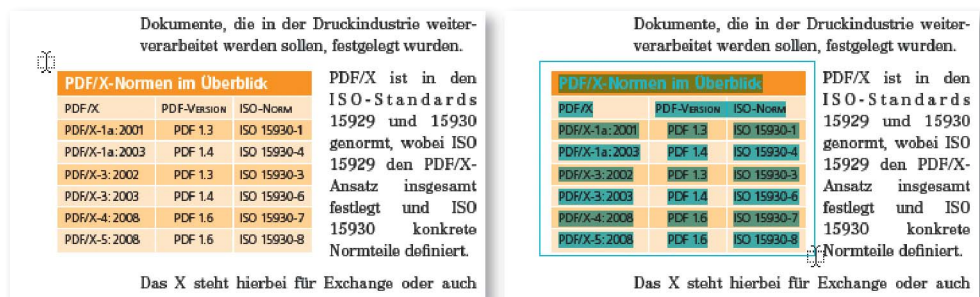
Wird Text mit dem »Auswahlwerkzeug für Text und Bilder« markiert, indem mit gedrückter Maustaste über den zu aktivierenden Text gefahren wird, werden häufig auch Bereiche markiert, die nicht benötigt werden.



Text mit gedrückter Alt-Taste markieren

Wenn die »Alt-Taste« gedrückt wird, ändert sich das Symbol der Einfügemarke. Nun kann der zu aktivierende Bereich einfach mit einem Auswahlrechteck markiert werden. Es werden dann nur Objekte aktiviert, die sich im Auswahlrechteck befinden.

Um nur einen bestimmten Bereich zu markieren, drücken Sie bei ausgewähltem AUSWAHLWERKZEUG FÜR TEXT UND BILDER die ALT-TASTE und ziehen ein Auswahlrechteck um den zu markierenden Bereich.

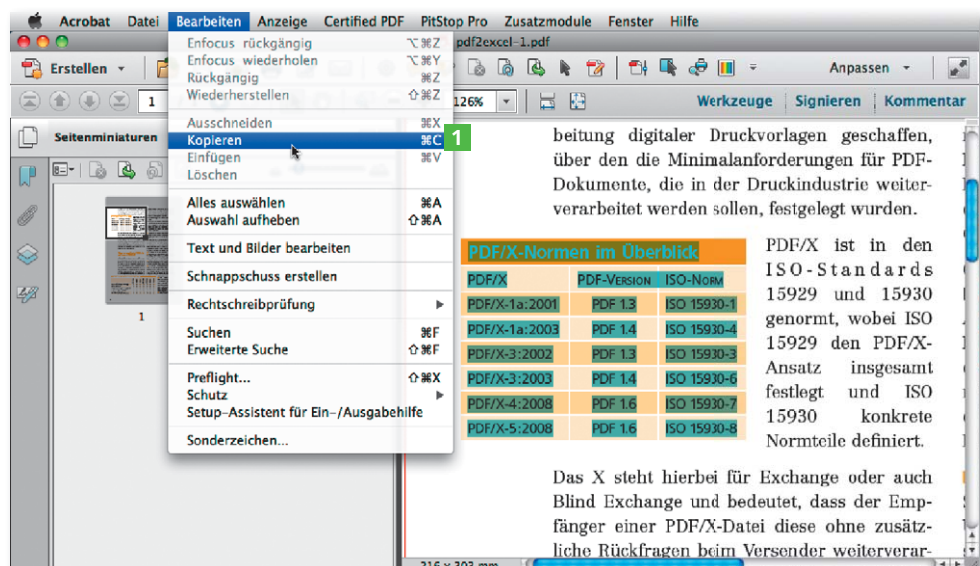


Texte können in einem PDF-Dokument kopiert und in ein anderes Programm eingefügt werden.

Übungsdatei

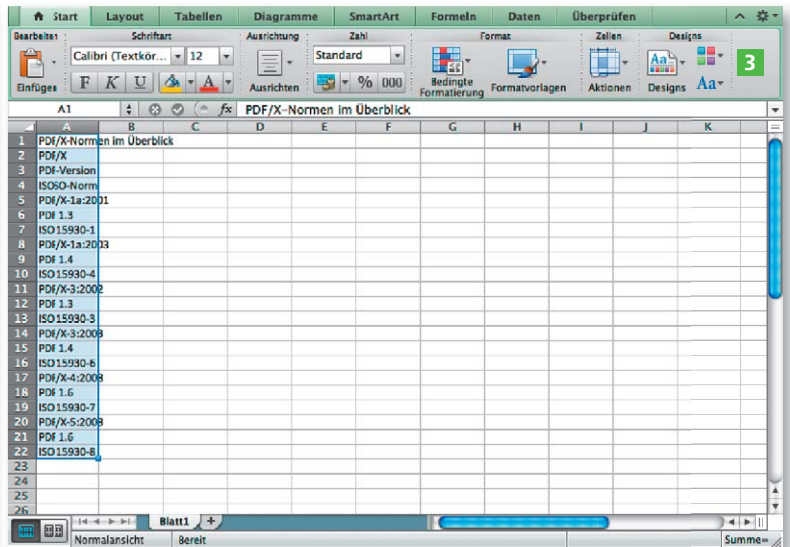
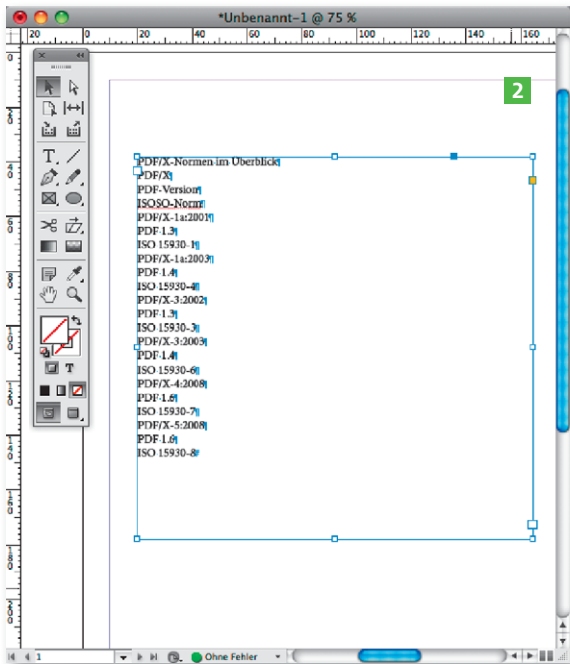
Die Beispieldateien finden Sie im Paket www.mev-medien.de/dateien/pdf/pdf26_pdf2excel.zip. Nach dem Entpacken liegen diese im Verzeichnis ARBEITSDATEIEN\ACROBAT\PDF2EXCEL.

Um den markierten Bereich in die Zwischenablage zu kopieren und dann in ein anderes Programm einzufügen, wählen Sie den Menübefehl BEARBEITEN → KOPIEREN 1.



Beim Einfügen gehen alle Formatierungen und die Tabellenanordnung verloren

Werden die kopierten Texte beispielsweise in ADOBE INDESIGN 2 in einen Textrahmen eingefügt, gehen dabei jegliche Formatierungsinformationen und die ursprüngliche Tabellenstruktur verloren. Jede Zeile wird mit einem Zeilenende abgeschlossen, sodass diese vor der eigentlichen Weiterverarbeitung erst einmal wieder gelöscht werden müssen. Wird der Inhalt in eine EXCEL-ARBEITSMAPPE 3 eingefügt, wird dort ebenfalls die Tabellenstruktur



Wird Text aus einem PDF-Dokument in ein anderes Programm eingefügt, werden die einzelnen Zeilen mit einer Zeilenschaltung abgeschlossen und die Originalformatierungen gehen beim Einfügen verloren.

ignoriert und jede Tabellenzelle wird in der ersten Spalte in eine separate Zelle eingefügt.

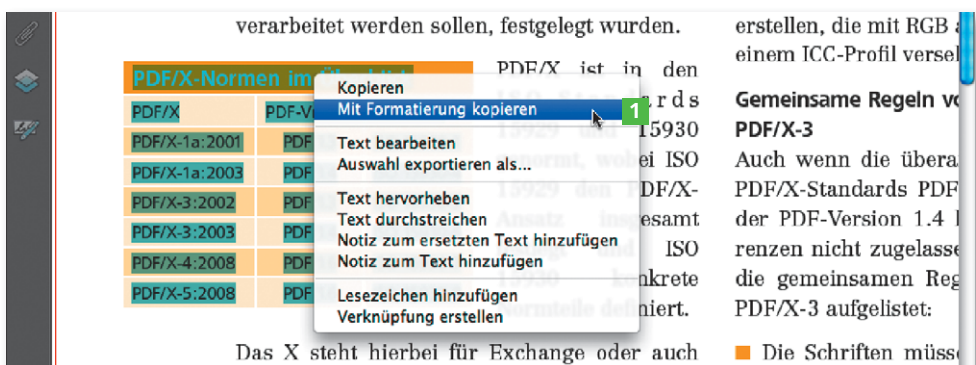
Text mit Formatierungen kopieren

Doch was viele Acrobat-Anwender nicht wissen, ist, dass ACROBAT XI PRO über einen Befehl verfügt, mit dem sie Texte mitsamt den Tabellenformatierungen kopieren und in ein Excel-Arbeitsblatt oder InDesign-Dokument einfügen können.

1. Dazu markieren Sie den zu kopierenden Text in Acrobat wie gewohnt mit dem AUSWAHLWERKZEUG FÜR TEXT UND BILDER. Anstatt des Menübefehls BEARBEITEN → KOPIEREN oder des entsprechenden Tastaturbefehls wählen Sie den Befehl MIT FORMATIERUNG KOPIEREN **1** aus dem Kontextmenü.

Hinweis

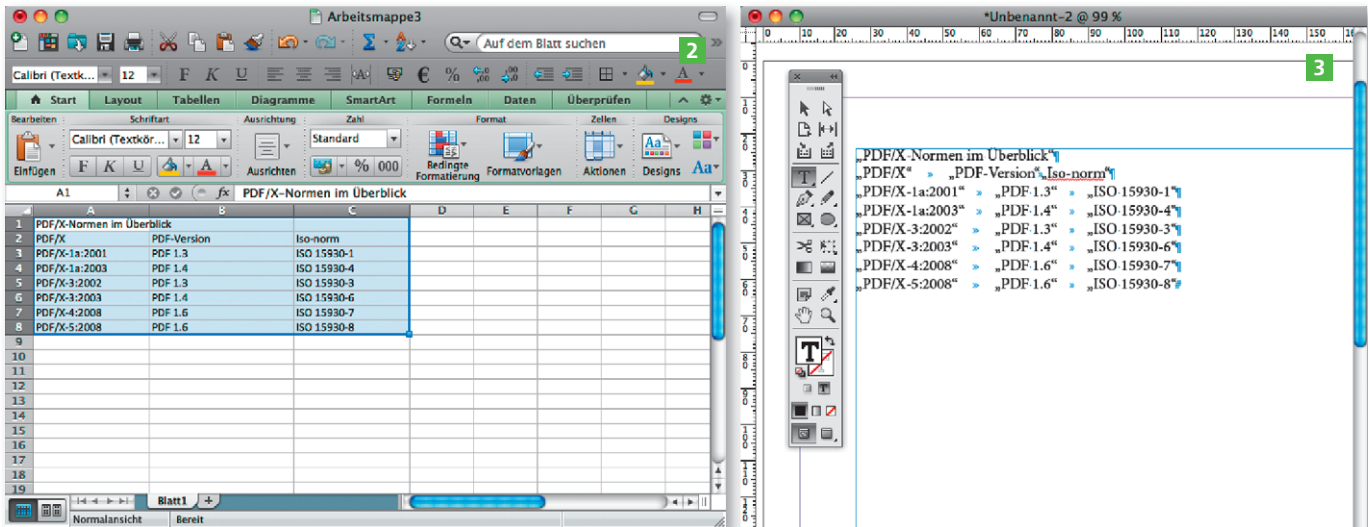
Das Kontextmenü erhalten Sie, indem Sie die rechte Maustaste drücken. Macintosh-Anwender, die mit einer Eintastenmaus arbeiten, drücken die CTRL-TASTE, um das Kontextmenü einzublenden.



Der Befehl zum Kopieren von markierten Texten/Tabellen kann über das Kontextmenü gewählt werden.

2. Wenn Sie den Text aus der Zwischenablage danach in eine EXCEL-ARBEITSMAPPE **2** einfügen, wird der Text mit der ursprünglichen Tabellenstruktur übernommen. Wird der auf diese Weise kopierte Text in ein INDESIGN-DOKUMENT **3** eingefügt, wird daraus zwar noch keine InDesign-Tabelle, doch werden die einzelnen Spalten nun durch einen Tabellensprung voneinander getrennt, sodass daraus über den entsprechenden Befehl sehr leicht eine Tabelle generiert werden kann. Mit diesem Befehl werden – anders als beim Einfügen in ein Microsoft-Word-Dokument –, werden die Schrifteninformati-

onen nicht kopiert und deshalb beim Einfügen in ein anderes Programm die jeweilige Standardschrift verwendet.



Wird Text über den Befehl »Mit Formatierung kopieren« aus einem PDF-Dokument in die Zwischenablage kopiert und anschließend in ein Microsoft-Word-Dokument eingefügt, werden die Formatierungen übernommen und der Text wird fortlaufend eingefügt.

Verwenden Sie den Befehl auch, wenn Sie eine Tabelle in einem PDF-Dokument kopieren und in QuarkXPress verwenden möchten, damit die Tabelle zeilenweise übernommen wird und die einzelnen Spalten durch einen Tabulator voneinander getrennt werden.


Ausgewählten Inhalt als Excel-Dokument speichern


Mit ACROBAT XI PRO besteht nicht nur die Möglichkeit, einzelne Bereiche zu markieren, zu kopieren und in eine Excel-Arbeitsmappe einzufügen. Acrobat bietet dem Anwender auch die Möglichkeit, einzelne Bereiche in einem PDF zu markieren und in ein Excel-Dokument zu konvertieren.


Wenn Sie in ACROBAT XI PRO Text auswählen und in ein Word-Dokument konvertieren möchten, kommt es darauf an, wie Sie die Inhalte im PDF-Dokument aktiviert haben.

Art der Textmarkierung

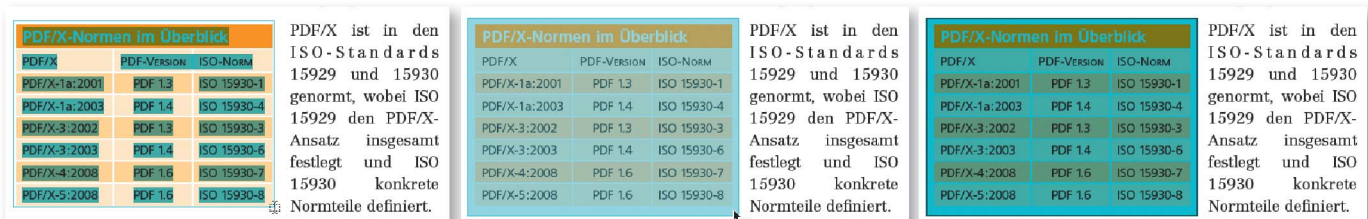
Erfolgt die Markierung durch das Markieren von Text, werden die einzelnen Texte nach dem Loslassen der Maustaste durch ein blau-transparentes Rechteck gekennzeichnet. Erfolgt die Markierung durch ein Auswahlrechteck, wird nach dem Loslassen der Maustaste der markierte Bereich durch ein blau-transparentes Rechteck gekennzeichnet.

1. Um einen beliebigen Text auf der Seite auszuwählen, wählen Sie das AUSWAHLWERKZEUG FÜR TEXT UND BILDER  aus der Werkzeugleiste und markieren den zu exportierenden Bereich.

■ Ziehen Sie mit dem AUSWAHLWERKZEUG FÜR TEXT UND BILDER  über den Bereich, den Sie exportieren möchten, und halten Sie dabei die ALT-TASTE gedrückt, damit Sie ein Auswahlrechteck erhalten.

■ Ziehen Sie mit dem AUSWAHLWERKZEUG FÜR TEXT UND BILDER  bei gleichzeitig gedrückter BEFEHLSTASTE (Macintosh) bzw. STRG-TASTE (Windows) ein Rechteck um die zu exportierenden Zeilen und Spalten auf.

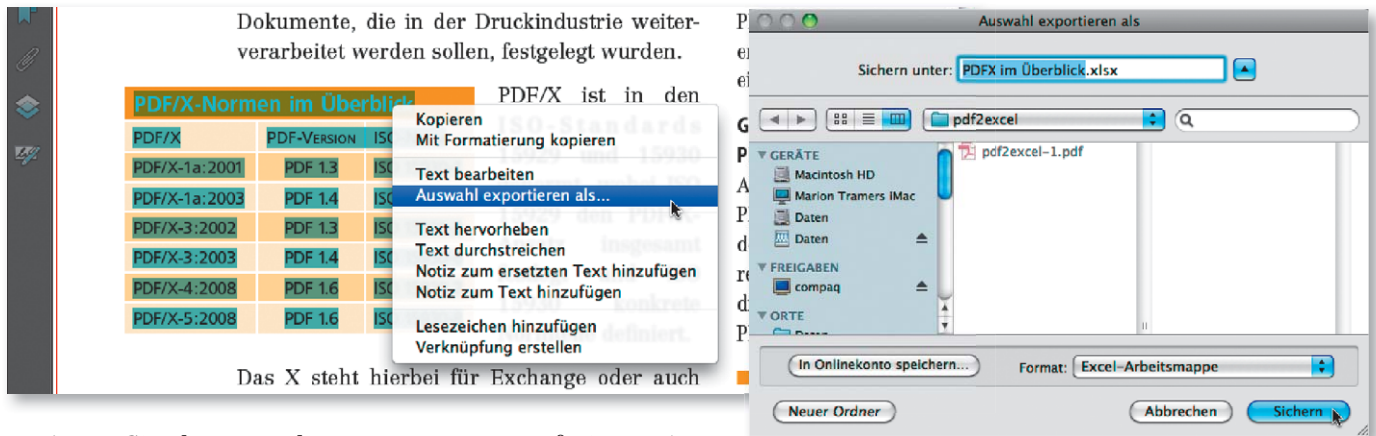
2. Nachdem Sie die Auswahl festgelegt haben, klicken Sie mit der Maus auf die Auswahl und wählen den entsprechenden Befehl aus dem Kontextmenü.



■ Wenn Sie das Kontextmenü auf einer Markierung öffnen, wählen Sie den Befehl AUSWAHL EXPORTIEREN ALS und im gleichnamigen Fenster den Speicherort sowie das Speicherformat aus und tragen den Dateinamen ein. Um die Auswahl in ein Excel-Dokument zu speichern, wählen Sie die Option EXCEL-ARBEITSMAPPE aus dem Einblendmenü FORMAT. Weitere Einstellungen zum Export können nicht definiert werden.

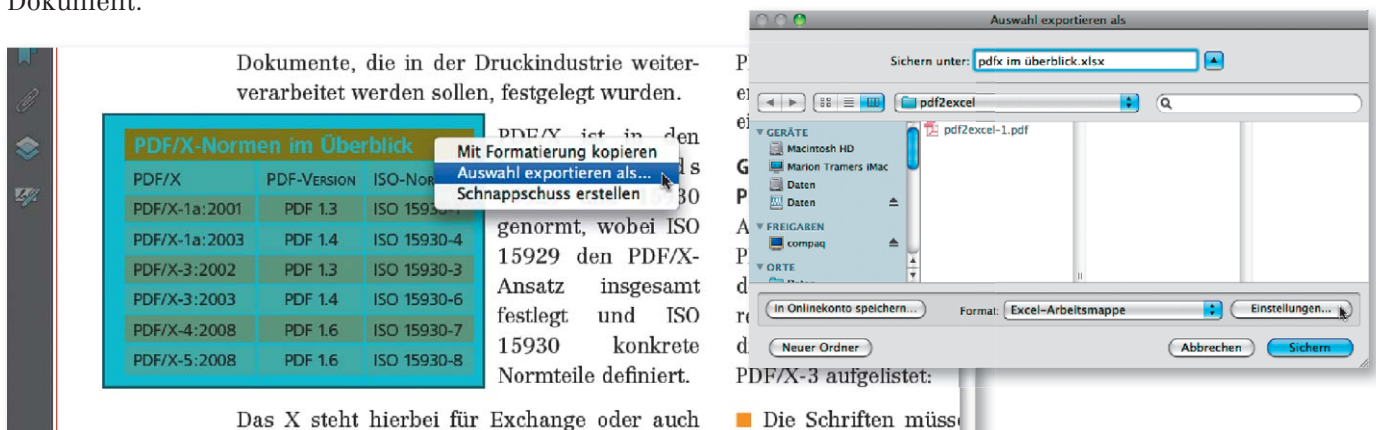
Kontextmenü auf Markierung

Wird das Kontextmenü auf einer Markierung geöffnet, enthält es mehrere Befehle. Über den Befehl »Auswahl exportieren als« kann die Markierung z. B. als Excel-Tabelle gespeichert werden.



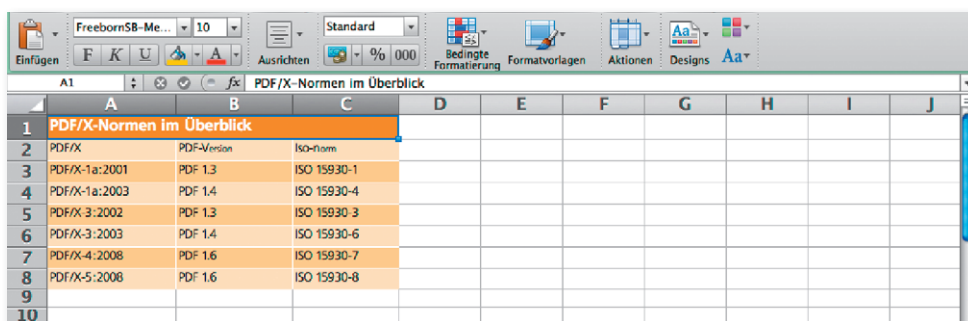
■ Wenn Sie dagegen das Kontextmenü auf einem Auswahlrechteck öffnen, enthält es lediglich drei Befehle zur Auswahl. Wählen Sie ebenfalls den Befehl AUSWAHL EXPORTIEREN ALS. Im gleichnamigen Fenster legen Sie den Speicherort und das Speicherformat fest und tragen den Dateinamen ein. Wenn Sie die Auswahl auf diese Weise exportieren, können Sie für das ausgewählte Speicherformat weitere Einstellungen vornehmen, indem Sie auf den Schalter EINSTELLUNGEN klicken. Die Einstellungen hier entsprechen den Einstellungen wie zum Speichern einer PDF-Datei als Word-Dokument.

Wird das Kontextmenü auf einem Auswahlrechteck geöffnet, enthält es nur drei Befehle. Über den Befehl »Auswahl exportieren als« kann die Markierung z. B. als Excel-Arbeitsmappe gespeichert werden.



3. Nach dem Export können Sie die Datei direkt in MICROSOFT EXCEL öffnen und das Ergebnis kontrollieren. Im Gegensatz zum Kopieren und Einfügen in

Die Tabelle wurde mitsamt der Text- und Zellenformatierung in die Excel-Arbeitsmappe exportiert.



ein Excel-Arbeitsblatt wurde beim Export die Text- und Zellenformatierung mit übernommen.

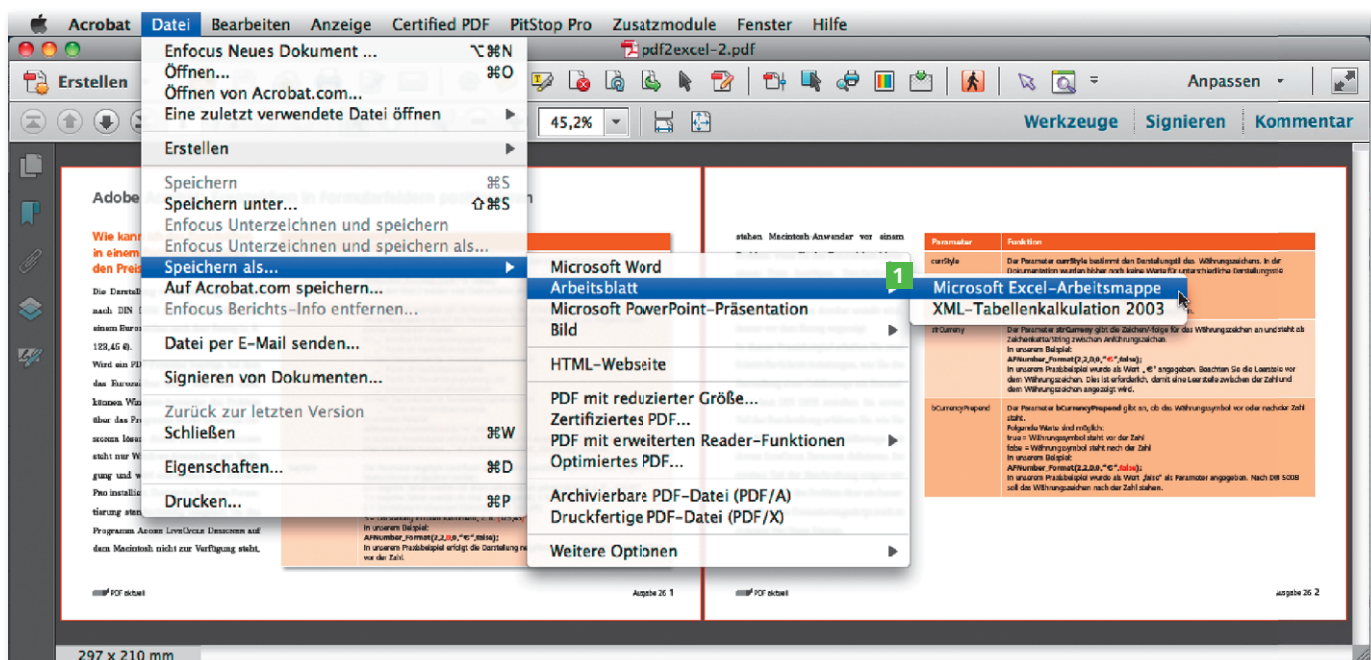
Wie kann ich ein PDF-Dokument in ein Excel-Dokument konvertieren?

Austausch mit Microsoft Excel

Über den Menübefehl DATEI → SPEICHERN ALS finden Sie im Untermenü ARBEITSBLATT den Befehl, mit dem Sie ein komplettes PDF als MICROSOFT EXCEL-ARBEITSMAPPE oder XML-TABELLENKALKULATION 2003 speichern können. Hierbei werden die Textformatierungen, wie z. B. Schriftart, -größe und -stil, so gut es eben geht mit in das Zieldokument übernommen. Enthält die PDF-Datei Fließtext, wird dieser zeilenweise in einzelne Tabellenzeilen gespeichert. Verwenden Sie die Funktion, um ein ganzes PDF-Dokument in ein Excel-Dokument zu konvertieren, wenn das PDF mehrseitige Tabellen beinhaltet.

Speichern als Excel-Dokument

1. Möchten Sie ein PDF-Dokument als MICROSOFT-EXCEL-DATEI speichern, wählen Sie den Menübefehl DATEI → SPEICHERN ALS → ARBEITSMAPPE **1** und den entsprechenden Befehl aus dem Untermenü. Wenn Sie das Aufgabenfenster WERKZEUGE einblenden, stehen Ihnen die Befehle zum Exportieren einer PDF-Datei als EXCEL-ARBEITSMAPPE oder XML-TABELLENKALKULATION 2003 auch über die Werkzeuggruppe INHALTSBEARBEITUNG über das Einblendmenü DATEI EXPORTIEREN NACH **2** zur Verfügung.



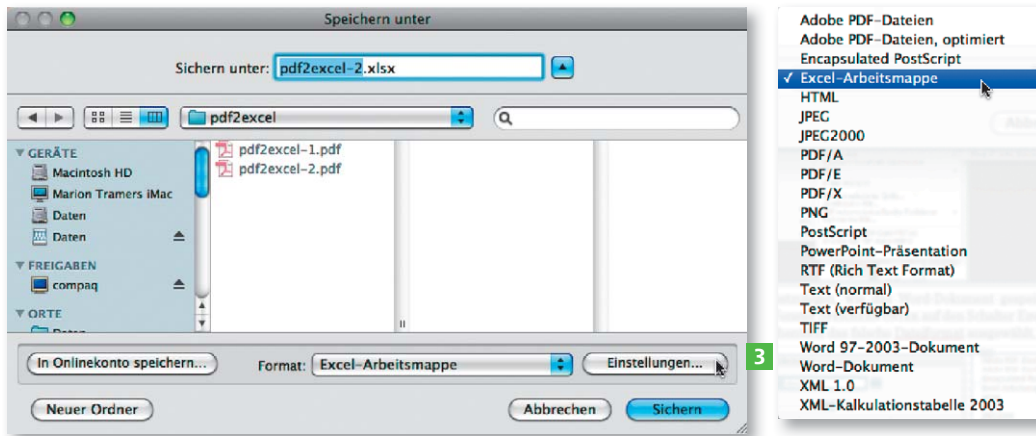
Über das Menü »Datei« stehen die Befehle zur Verfügung, mit denen ein PDF-Dokument als Microsoft-Excel-Dokument gespeichert werden kann.

■ Abhängig von der Excel-Version, mit der die Datei später weiterverarbeitet werden soll, wählen Sie im Untermenü die Option MICROSOFT EXCEL-ARBEITSMAPPE, wenn Sie ein XLSX-DOKUMENT erstellen möchten, das nur mit einer

MICROSOFT-OFFICE-EXCEL-VERSION 10 oder höher geöffnet werden kann. Soll das Dokument auch in einer früheren Version von MICROSOFT EXCEL zu öffnen sein, wählen Sie die Option XML-TABELLENKALKULATION 2003.

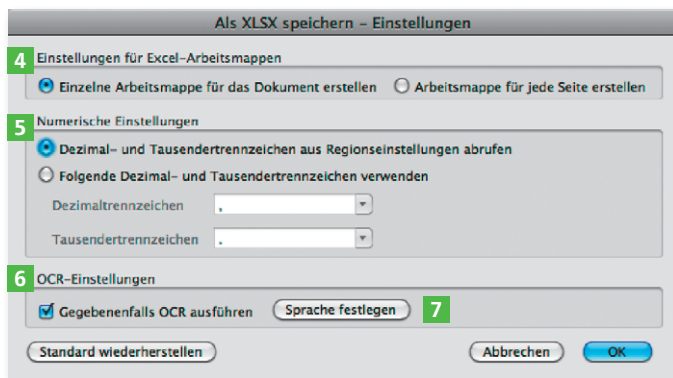
2. Um die Einstellungen festzulegen, wie das Excel-Dokument gespeichert werden soll, klicken Sie im Fenster SPEICHERN UNTER auf den Schalter EINSTELLUNGEN 3.

Hinweis
Haben Sie beim SPEICHERN ALS versehentlich das falsche Dateiformat ausgewählt, können Sie über das Einblendmenü FORMAT das Speicherformat nachträglich ändern.



Im Fenster »Speichern unter« kann das Dateiformat nachträglich geändert werden. Um die Speichereinstellungen festzulegen, klicken Sie auf den Schalter »Einstellungen«.

3. Die Einstellungen zum Speichern als MICROSOFT EXCEL-ARBEITSMAPPE und als XML-TABELLENKALKULATION 2003 sind identisch. Die resultierende Excel-Datei unterscheidet sich lediglich im Dateiformat und in der Version, mit der das exportierte Dokument später geöffnet werden kann.

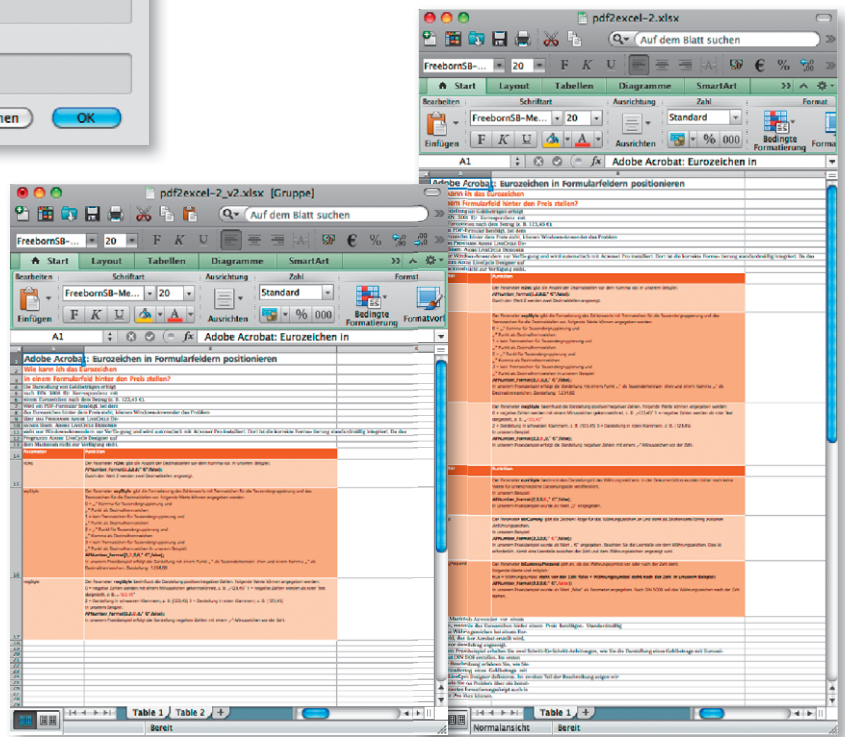


Die Einstellungen zum Speichern als XLSX-Dokument (links) oder als XML-Kalkulationstabelle (rechts) sind identisch.

Links: jede Seite eine Arbeitsmappe; rechts: Seiten wurden fortlaufend hintereinander angehängt

■ Im Bereich EINSTELLUNGEN FÜR EXCEL-ARBEITSMAPPEN 4 aktivieren Sie die Option EINZELNE ARBEITSMAPPE FÜR DAS DOKUMENT ERSTELLEN, wenn die einzelnen Seiten des PDF-Dokuments fortlaufend in einer Arbeitsmappe angehängt werden sollen. Wenn Sie die Option ARBEITSMAPPE FÜR JEDE SEITE ERSTELLEN aktivieren, wird für jede Seite im PDF eine neue Arbeitsmappe im Excel-Dokument erstellt.

■ Im Bereich NUMERISCHE EINSTELLUNGEN 5 legen Sie das Format für Dezimal- und Tausendertrennzeichen für numerische Daten fest. Sie können die

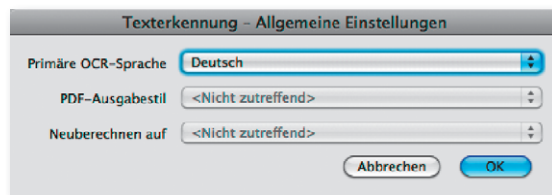


DEZIMAL- UND TAUSENDERTRENNZEICHEN AUS REGIONSEINSTELLUNGEN ABRUFEN, die auf Ihrer Arbeitsstation gespeichert sind, oder die Trennzeichen in den entsprechenden Feldern eintragen, wenn Sie die Option FOLGENDE DEZIMAL- UND TRENNZEICHEN VERWENDEN aktivieren.

OCR ausführen

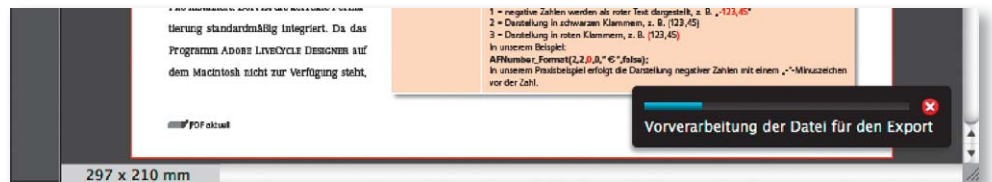
■ Enthält ein PDF-Dokument gescannte Texte, können diese als Bild oder als bearbeitbare Texte gespeichert werden. Wenn Sie die Texte beim Speichern als Excel-Dokument gleichzeitig in bearbeitbaren Text konvertieren möchten, aktivieren Sie die Option GEBEBENENFALLS OCR AUSFÜHREN im Bereich OCR-EINSTELLUNGEN **6**. Um die OCR-Sprache auszuwählen, klicken Sie auf den Schalter SPRACHE FESTLEGEN **7**.

Wird gescannter Text gleichzeitig in bearbeitbaren Text umgewandelt, muss die Sprache für die Umwandlung festgelegt werden.



4. Um das Dokument zu speichern, klicken Sie auf den Schalter SICHERN. Im unteren rechten Bereich des Acrobat-Fensters wird daraufhin ein Verarbeitungsbalken angezeigt, anhand dessen Sie die Verarbeitung des Dokuments erkennen können.

Anhand des Verarbeitungsbalkens kann die Speicherung des PDF als Word-Dokument mitverfolgt werden.



5. Danach können Sie das Dokument direkt in MICROSOFT EXCEL **8** öffnen oder auch in ein anderes Programm importieren.

Nach dem erfolgreichen Export kann die Tabelle direkt in Excel geöffnet und bearbeitet werden. Die Daten können auch in andere Programme, wie z. B. Adobe InDesign, importiert werden.

6. Wenn Sie die Tabelle beispielsweise in ADOBE INDESIGN **9** laden, können Sie festlegen, dass die gesamte Tabellenformatierung geladen wird. Mit ein wenig Aufwand kann die Tabelle dort dann entsprechend aufbereitet werden.

