

Die Grundelemente des World Wide Web sind:

? HTTP(HyperText Transfer Protocol) - der Datenaustausch-Standard der von Computern benutzt wird, um miteinander zu kommunizieren und Dateien zu teilen.

? URL's (Uniform Resource Locator) - die Adresse von Dateien oder Verzeichnissen auf dem Web.

? HTML (HyperText Markup Language) - die Programmier-"Tags", die Textdokumenten angefügt werden, und diese dann in Hypertext-Dokumente verwandeln.

HTTPHTT steht für HyperText Transfer Protocol -- die Methode mit der Hypertext-Dateien über das Internet ausgetauscht werden. Lassen Sie uns die Ausdrücke einzeln betrachten. "Hypertext" wurde 1965 von Ted Nelson geprägt und bedeutet, daß Text nicht linear angeordnet sein muß. Auf dem Web ist dies der Text der mit einer anderen Datei verbunden ist. Wenn Sie auf ein Wort klicken, und zu einer anderen Seite (oder einer Sounddatei oder einem Bild) gelangen, dann verwenden Sie Hypertext. World Wide Web Seiten, die in HTML geschrieben sind, verwenden Hypertext, um zu anderen Dokumenten zu springen.

"Hypertext Transfer" ist das Übertragen von Hypertext-Dateien von einem Computer zu einem anderen. Wenn Sie ein Hypertext-Dokument lesen, können Sie auf ein Link klicken, und begeben sich zu einer anderen Site. Dieses Dokument wurde dann von einem anderen Computer zu Ihrem Computer über das Internet übertragen.

Und was bedeutet nun "Protocol"? In der Computerwelt ist ein Protokoll eine Sammlung von Regeln und Konventionen, die die Übermittlungsart zwischen zwei Computern beschreiben. Zusammengefaßt ist ein HyperText Transfer Protocol eine Konvention, die verwendet wird, um den Austausch von Hypertext-Dateien (Webseiten) auf dem Internet möglich zu machen.

Hypertext

Eine Art Informationen darzustellen, in denen Text, Sound, Grafiken und Bilder miteinander verbunden sind, damit der Benutzer in keiner bestimmten Reihenfolge zwischen diesen Dateien umherspringen kann. Hypertext bezieht sich im allgemeinen auf das World Wide Web, das viele Links zu anderen Dokumenten enthält.

Link

Bezieht sich im allgemeinen auf hervorgehobene Wörter oder Satzteile in einem Hypertext Dokument. Das Link erlaubt es Ihnen, von einem Bereich desselben Dokuments oder zu einem anderen Dokument im World Wide Web zu "springen".

HTML

HTML ist ein Akronym für HyperText Markup Language. HTML ist die Computersprache für die Erstellung von Hypertext Dokumenten. HTML nutzt eine bestimmte Liste von Tags, die die allgemeine Struktur von verschiedenen Dokumenten, die auf dem World Wide Web miteinander verbunden sind, beschreiben.

Tags

Tags sind die Formatierungselemente, die in HTML-Dokumenten verwendet werden. Sie weisen den Web-Browser an, wie Text und Grafiken auf einer Webseite dargestellt werden sollen. Zum Beispiel: Um Text in Kursiv darzustellen, wird das "i" Tag am Anfang und Ende des Texts verwendet.

Dateiformate und -konvertierung

Welche Dateiformate sind für Programm-übergreifenden Datenaustausch empfehlenswert? Und wie kann man dem Empfänger einer Datei etwas helfen? Hier ein paar Tipps dazu.

Info an Computer-Neulinge: Exportieren in ein Dateiformat bedeutet nicht, einfach die Datei-Endung zu ändern. Für den Datei-Export muss das Dokument durch den entsprechenden Export-Filter! Diesen erhält man z.B. in Winword mit "Datei" -> "Speichern unter", und dann den "Dateityp" auswählen.

Das Problem

Nicht jeder benutzt das selbe Grafik-, Text- oder Datenbank-Programm. Und wenn, dann will's der Teufel, dass es nicht die selbe Programm-Version ist, die Freunde oder Geschäftspartner verwenden, oder etwa nicht einmal auf dem selben Betriebssystem. In den meisten Fällen gibt es jedoch

Möglichkeiten, den Inhalt einer Datei auch BenutzerInnen anderer Programme oder Betriebssysteme zugänglich zu machen.

Dazu gibt es zwar in fast jedem Programm Export- oder Import-Möglichkeiten. Aber auch wenn das Dateiformat unterstützt wird, kann es sein, dass man gewisse Verluste in der Gestaltung in Kauf nehmen muss.

Ein paar Standards oder Pseudo-Standards haben sich zwar entwickelt, was nicht zuletzt auch ein grosser Verdienst des World Wide Web ist. Aber nicht einmal HTML, die Dokumentensprache des Internet, besteht den Standard-Test, denn die meisten Webseiten, die man mit drei verschiedenen Browsern liest, sehen auch dreimal verschieden aus - oft nur geringfügig, manchmal jedoch auch total! Fast jeder Software-Hersteller betrachtet seinen eigenen Weg als der absolut richtige. Es gibt sogar Dateiformate, die praktisch nur von Programmen aus dem eigenen Haus gelesen werden können.

#### Mit Verlust rechnen

Wer Dateien zwischen verschiedenen Betriebssystemen oder Programmen austauschen will, muss sich damit abfinden, dass am anderen Ende praktisch nie genau das herauskommt, was man am einen Ende verschickt hat.

Als hilfreich entpuppen sich jeweils die Export- und Import-Optionen, die fast jedes Programm mitbringt. Oft gehen in Text-Dokumenten Formatierungen, wie Einzüge, Ränder, Aussparungen, Tabulatoren etc. verloren.

Je komplizierter die Formatierung, desto unwahrscheinlicher wird es sein, dass jemand, der das Dokument in eine andere Applikation importiert, diese Daten gleich verwenden kann.

#### Lesergruppe berücksichtigen

Wenn die Datei zwischen verschiedenen Microsoft Office-Applikationen und -Versionen ausgetauscht wird, kann man einigermassen beruhigt aufatmen. Die meisten Office-Applikationen beinhalten Filter, welche ein Speichern in einer früheren Version ermöglichen.

Eine Ausnahme bildet MS Access 97. Wer darauf eine Datenbank erstellt, wird Mühe haben, diese verlustfrei in Access 95 oder gar Access 2.0 zu exportieren.

Soll die Datei nur angezeigt und/oder ausgedruckt werden können? Auch das ist mit Dateien aus Office 97 kein Problem. Denn für Word, Excel und Powerpoint 97 existieren sogenannte "Viewer", das sind Anzeigeprogramme, welche man sich von der Microsoft Homepage (und neu auch hier) herunterladen kann. Die Dateien können damit allerdings nicht bearbeitet werden.

Auf manchen Homepages und auf den CD-ROMs vieler Software-Hersteller finden sich Dokumente, die im PDF-Format gespeichert sind. Diese Dateien lassen sich ausschliesslich mit dem extrem teuren Adobe Acrobat bearbeiten und erstellen. Dafür ist aber selbstverständlich wenigstens ein Gratis-Viewer verfügbar, nämlich der Adobe Acrobat Reader (derzeit in der Version 4.0).

Mit einem - zugegebenermassen relativ aufwändigen - Trick geht das Speichern in PDF jedoch auch anders. Mehr darüber im Sonderbeitrag PDF-Info.

#### Textformate

Die neueren Versionen von professionellen Textverarbeitungs-Programmen enthalten für den Datei-Import fast immer die viel besseren Filter als für den Datei-Export. Der Grund dafür liegt im Bestreben der einzelnen Hersteller, den Umstieg von der fremden Applikation auf die eigene zu erleichtern - und natürlich nicht umgekehrt.

Wer Daten zwischen Benutzern verschiedener Text-Programme austauschen möchte, sollte zunächst versuchen, Texte in die nächst-ältere Version des Programms zu exportieren, das man selber verwendet. Beispiel: Wer Word 97 hat, kann die Dateien im Word 6.0 oder 7.0 Format speichern. Für dieses wird das Programm des Empfängers möglicherweise bereits einen Import-Filter besitzen. Für Mac existieren übrigens manche Microsoft Office Programme ebenfalls. Beispielsweise zwischen MS Word für Mac und MS Word für Windows können Dateien in der Regel ohne grössere Verluste ausgetauscht werden.

Fast jedes Text-Programm besitzt eine Möglichkeit, Dateien als RTF zu speichern und zu öffnen. RTF (Rich Text Format) unterstützt grundlegende Formatierungen wie Zeilenabstände, Tabulatoren, Schriftarten. Dies wird meist reichen, den Inhalt einer Datei an die Leute zu bringen. Mac-BenutzerInnen sollten daran denken, dass die Dateinamen nach dem Export in's RTF auch die Endung ".RTF" tragen sollten, da Windows den Dateityp anhand der Endung identifiziert.

Andere, wie StarOffice und MS Word (ab Version 97) können Dateien auch im HTML-Format speichern. Einzüge, Tabulatoren etc. gehen dabei verloren, Schriftarten und -größen werden dabei jedoch beibehalten.

Export als ASCII-Text ("nur Text") kann als letzte Möglichkeit dienen. Dabei werden alle Formatierungen entfernt. Es werden nur die Textzeichen übernommen, ohne Schriftarten und dergleichen. Im Zweifelsfall sollte man in Betracht ziehen, Dateien in verschiedenen Formaten anzubieten.

## Bildformate - die gängigen

Jedes Grafik-Programm bringt eigene Datei-Formate mit, wie der Micrografix Picture Publisher mit PPF, Adobe Illustrator mit AI und noch beliebige mehr. Als weltweite, programm-unabhängige Standards gelten heute hauptsächlich TIFF (Tagged Image File Format), GIF (CompuServe Graphics Image Format) und JPEG.

Im professionellen Werbe- und Druckbereich wird für die Aufbereitung von Prospekten oft TIFF verwendet. Solche Dateien umfassen aber oft mehrere Megabytes, denn sie beinhalten eine extrem hohe Auflösung und Farbtreue.

Im Normalfall reicht es, Bilder als JPG oder GIF zu speichern. Beide Formate haben sich als Standard im Internet hervorgetan, lassen auch eine relativ hohe Bildqualität mit einer verhältnismäßig kleinen Dateigröße zu und jeder, der einen Webbrowser hat, kann diese Dateien anzeigen. Welches Format sollte man aber nun wählen? Das ist ziemlich einfach.

Für Bilder mit wenigen Farben und ohne Farbverläufe wählt man GIF. Das könnten Logos oder Schriftzüge sein, Icons, einfache Strichzeichnungen oder Buttons auf einer Webseite. Wenn es sich um Fotografien handelt, wird man eher JPG wählen. Hier ist auch die Kompression zu beachten. Je höher die Kompression, desto schlechter die Bildqualität und Detailtreue. Je geringer die Kompression, desto größer die Datei. Im Normalfall wird man eine Kopie der Originaldatei machen und damit etwas herumexperimentieren, bis man das beste Verhältnis zwischen Dateigröße und Bildqualität hat.

## Windows Bitmap (BMP)

Auch bekannt als: BMP, DIB, Windows BMP, Windows DIB, Compatible Bitmap

### Hintergrund

Das Bitmap-Format stammt von Microsoft und wird unter Windows vom System zur Darstellung von Bildern (z.B. Hintergrundbilder), sowie von einigen Programmen, wie z.B. dem mitgelieferten MS-Paint, benutzt.

Von diesem Format gibt es verschiedene Abarten und auch teilweise kompatible OS/2-Versionen. Diese Vielfalt teils unterschiedlicher Implementierungen des selben Formats lässt es im Vergleich zu anderen besser spezifizierten Formaten in keinem guten Licht erscheinen.

### Kompression

Es wird RLE-Kompression unterstützt, die aber (außer in Ausnahmefällen) nicht sehr effektiv ist.

### Eignung

Ein Vorteil des Formates ist, daß es ungeheuer viele Programme auch außerhalb der Windowswelt unterstützen. Problematisch ist daran andererseits wieder, daß es sehr viele unterschiedliche Versionen des Formates gibt - von denen wiederum nicht jedes Programm alle versteht. Gerade die OS/2-Abarten des BMP-Formates werden von Windowsprogrammen oft nicht verstanden, wenn es auch umgekehrt funktioniert. Für komplexe Bilder ist das Format wegen der unzureichenden Kompression ohnehin eher schlecht geeignet.

Fazit: Das Format hat zwar keine gravierenden Nachteile; im allgemeinen sind andere Formate (wie zum Beispiel JPEG oder TIFF) aber vorzuziehen.

### Erkennungsmerkmale in der Datei

Jede BMP-Datei beginnt mit BM. Das sind die ersten beiden Buchstaben in der Datei, sie ist dadurch eindeutig erkenntlich.

## Encapsulated PostScript

Auch bekannt als: EPS, EPSF, EPSI

#### Hintergrund

Encapsulated PostScript ist eine Abart von PostScript. Es führt zusätzlich verpflichtende Kommentare und eine Dokumentstruktur ein, die es einem Programm erleichtern, die Datei zu lesen.

#### Kompression

EPS bietet keine Kompression.

#### Eignung

Es ist sehr einfach, EPS-Dateien in PostScript- oder TeX-Dokumente einzubinden. Außerdem wird das Format von sehr vielen vektororientierten Zeichenprogrammen (wie zum Beispiel Corel Draw) unterstützt. Für die Speicherung von Pixelbildern ist es ungeeignet. Das ist zwar möglich, aber nicht sinnvoll.

#### Erkennungsmerkmale in der Datei

Wie auch PostScript ist eine EPS-Datei menschenlesbar. Man erkennt sie daran, daß in der ersten Zeile die Buchstabenfolge EPSF (für Encapsulated PostScript File) vorkommt.

#### GIF

(Graphics Interchange Format)

#### Hintergrund

Dieses Format wurde ursprünglich von der Firma CompuServe entwickelt, um Bilder für die Onlinebetrachtung zur Verfügung zu stellen. Es ist heute das wohl am meisten verbreitete Format zur Darstellung von Pixelbildern bis 256 Farben.

#### Kompression

GIF kann mit LZW-Kompression arbeiten, die gerade bei großen gleichfarbigen Flächen sehr effizient ist.

#### Eignung

GIF ist vor allem für kleine Buttons und Bildchen auf Webseiten sehr gut geeignet. Es kann zwar maximal 256 Farben aufnehmen; für diesen Zweck reicht das aber auf jeden Fall aus. Nett ist außerdem, daß man eine transparente Farbe definieren kann, so daß der Hintergrund der Webseite sichtbar wird. Für kleinere Animationen gibt es die Möglichkeit, mehrere Bilder in einer Datei abzuspeichern, die dann vom Browser nacheinander abgespielt werden (Animated GIF). Ungeeignet ist das Format für Fotos. Zum einen ist hier die maximale Farbtiefe von 8 bit (256 Farben) unbefriedigend; zum anderen ist der Kompressionsalgorithmus bei Farbverläufen und ähnlichem, wie sie bei Fotos oft auftreten, nicht besonders effizient.

#### Erkennungsmerkmale in der Datei

Jede GIF-Datei beginnt mit den Buchstaben GIF. Dadurch ist sie eindeutig zu identifizieren.

#### JPEG

(JPEG File Interchange Format, JFI, JPG, JPEG)

#### Hintergrund

JPEG (Joint Photographic Experts Group) ist an sich kein Dateiformat, sondern eine Form der Kompression, die speziell für digitalisierte Fotos entwickelt wurde. Deshalb müßte das Format auch korrekt JFIF heißen, also JPEG File Interchange Format. Es hat sich aber inzwischen der Begriff JPEG für das Datenformat ebenso durchgesetzt.

#### Kompression

Hier handelt es sich um die eben erwähnte JPEG-Kompression. Sie ist sehr effektiv und kann selbst sehr große Bilder mit kaum merklichem Qualitätsverlust stark verkleinern. Aber da sind wir schon beim Thema: Es handelt sich um verlustbehaftete Kompression; das Bild wird also beim komprimieren "schlechter", wenn auch (je nach Stärke der Kompression) meist unmerklich. Das schöne ist, daß man den Qualitätsverlust selbst bestimmen kann: Bei der Speicherung einer JPEG-Datei kann man den Kompressionsfaktor einstellen, der sich dann in der Qualität des resultierenden Bildes widerspiegelt.

#### Eignung

JFIF ist das ideale Format, um Fotos im Internet darzustellen. Je nach Kompressionsfaktor können die Dateien unglaublich klein werden, und das Bild ist trotzdem noch schön anzusehen.

Achtung: JPEG ist völlig ungeeignet zur Kompression von Strichzeichnungen, Pixelbildern, die Schrift enthalten, und ähnlichem! Scharfe Ränder verwischen dabei nämlich. Bei einem Foto ist das meist nicht weiter wichtig oder fällt sogar gar nicht auf.

## Portable Document Format (PDF)

Drucken unter Unix: Adobe Acrobat Reader

Drucken unter Windows: Adobe Acrobat Reader

### Hintergrund

Kompression

Eignung

PDF wird für Handbücher, Datenblätter und Ähnliches verwendet. Es ist wie PostScript plattformunabhängig, und es existieren Acrobat-Reader von Adobe für sämtliche Plattformen (Mac, Windows, Linux, SUN,...).

Downloaden können Sie den Acrobat Reader direkt bei Adobe-Produkte , oder für Windows bei den Windows-Tools.

Aber auch z.B. Corel-Draw kann PDF lesen.

## Portable Network Graphic Format (PNG)

### Hintergrund

PNG (sprich: Ping) ist ein sehr neues Bildformat, das entwickelt wurde, um GIF zu ersetzen. Das zum einen deshalb, weil GIF nur 256 Farben unterstützt, zum anderen (wichtiger), weil GIF die LZW-Kompression benutzt, die nicht frei verfügbar ist. PNG steht auch für "PNG is not GIF". Es unterstützt Farbtiefen von bis zu 48 Bit (!) und benutzt einen recht effektiven, verlustfreien Kompressionsalgorithmus.

Bisher wird das Format leider noch von relativ wenigen Programmen vollständig unterstützt. Da aber auch das W3C (das Konsortium, das HTML-Standards verabschiedet) die Benutzung von PNG im WWW empfiehlt, ist davon auszugehen, das sich das Format in den nächsten Jahren schnell verbreiten wird.

Kompression

Eine Variante von LZ77, die recht effektiv und zusätzlich verlustfrei ist.

Eignung

PNG ist vor allem für Webseiten gedacht. Aber auch für die Speicherung ganz "normaler" Bilder stellt es eine interessante Alternative dar, da die Dateien recht klein werden und dabei kein Qualitätsverlust entsteht.

Erkennungsmerkmale in der Datei

## Tag Image File Format (TIFF)

### Hintergrund

Die erste Version von TIFF wurde 1986 von der Aldus Corporation entwickelt. Weithin genutzt wird das Format seit der Version 4.0. Die momentan aktuelle Version ist 6.0; viele Bilder liegen aber noch in den Versionen 4.0 und vor allem 5.0 vor.

TIFF ist ein sehr komplexes Format, das ungeheuer viele Möglichkeiten der Speicherung von Bitmapdaten bietet. Es ist gewissermaßen die eierlegende Wollmilchsau unter den Pixelformaten. Das ist sein Vorteil und sein Nachteil: Zum einen kann das Format sehr vielseitig genutzt werden; zum anderen verstehen aber einige Programme nicht jede Seltsamkeit, die in einer TIFF-Datei angewandt werden kann. Bei professionellen Programmen kann man aber davon ausgehen, daß die Spezifikation vollständig erfüllt ist.

Kompression

RLE, LZW, CCITT (Group 3 und Group 4), JPEG und unkomprimiert.

Eignung

TIFF ist hervorragend geeignet zur Speicherung größerer Bilder ohne Qualitätsverlust. Es ist als das professionelle Format in der Bildbearbeitung zu betrachten.