

## Management Summary zur KOST-Studie: *PDF/A-2 und PDF/A-3: Was ist neu?*

### 1 Einleitung

Das Summary fasst die Studie zusammen, welche im Rahmen der Arbeiten an Version 3 des *Katalogs archivischer Dateiformate KaD* der KOST entstand. Diese Studie dokumentiert die Veränderungen zwischen den PDF/A-Versionen 1, 2 und 3. Die Veränderungen gehen einerseits auf die verschiedenen Referenzversionen zurück, andererseits auf die unterschiedlichen Einschränkungen gegenüber der jeweiligen Referenz.

### 2 Gesamtbeurteilung

#### PDF/A-2

PDF/A-2 erlaubt neu die Verwendung der Transparenz sowie der JPEG2000-Komprimierung. Dass die Transparenz jetzt erlaubt ist, ist grundsätzlich von grossem Vorteil, da besonders in diesem Bereich bei der Konvertierung in PDF/A-1 gewisse Mängel existieren. Durch die JPEG2000-Komprimierung kann im Vergleich zur Dateigrösse eine bessere Qualität erzielt werden. Diese beiden Eigenschaften sind sehr positiv zu beurteilen und auch nicht bestritten.

Ein gewisses Umdenken ist erforderlich bei der Verwendung der neu erlaubten Eigenschaften *Embedded Files* und *Optional Content*, da sich damit die eindeutige visuelle Reproduzierbarkeit nur noch auf die Erscheinung am Computer bezieht.

Kritisch muss die Eigenschaft *Collections* beurteilt werden. Dabei empfiehlt die KOST, die *Portfolios* ohne eine inhaltlich wichtige Struktur zu erstellen, da gewisse Reader diese Struktur nicht darstellen können.

**PDF/A-2 ist aus der Sicht der KOST ein Format, das sich für die dauerhafte Archivierung eignet. Es ist allerdings im Gegensatz zu PDF/A-1 nicht mehr zu 100% als "Druckformat" geeignet, sondern ein Format, das digital bleiben muss.**

**Bei der Verwendung von *Collections* wird empfohlen, die *Portfolios* ohne eine inhaltliche wichtige Struktur zu erstellen, da diese nicht unbedingt vermittelt werden kann.**

#### PDF/A-3

PDF/A-3 ist mit einer wesentlichen Ausnahme identisch mit PDF/A-2. Die Ausnahme betrifft die *Embedded Files*. Im Gegensatz zu PDF/A-2, bei welchem nur *Embedded Files* erlaubt sind, die selber PDF/A-1- oder PDF/A-2-konform sind, ist in PDF/A-3 keine solche Einschränkung mehr vorhanden.

Dies bedeutet, dass in einer *Collection* oder als *Attachment* irgendwelche Dateien angehängt werden dürfen, von einem Word-Dokument über einen Film bis hin zu einem Programm (EXE). Entsprechend hat das Archiv keine Kontrolle mehr über die Anzahl archivierter Formate und kompromittiert dadurch das *Preservation Planning*. *Attachments*, welche nicht

PDF-Dateien sind, werden zudem auch nicht im PDF-Viewer geöffnet, sondern brauchen die entsprechende Spezialapplikation. Dies bedeutet, dass das Migrationsprinzip nicht mehr oder nur noch mit sehr grossem Aufwand möglich ist und die Datei möglicherweise nur via Emulation noch gelesen werden kann.

**PDF/A-3 ist aus der Sicht der KOST ein Format, das sich nicht für die dauerhafte Archivierung eignet. Entsprechend sollten öffentliche Archive auf PDF/A-3 verzichten. Für andere Archive wie z.B. Firmenarchive mit beschränkter Archivierungsdauer kann es jedoch interessant sein.**

### 3 Glossar

#### ***Optional Content***

*Optional Content* ermöglicht u.a. das Ein- und Ausblenden diverser Ebenen / Layer. Damit kann das gleiche Dokument eine komplett andere Erscheinung annehmen. Diese geschieht jedoch bereits im Original-PDF (ab Version 1.5).

*Optional Content* wird häufig bei CAD-Zeichnungen, Plänen und mehrsprachigen Dokumenten verwendet.

#### ***Embedded Files***

Eine PDF-Datei kann *Embedded Files* enthalten. Diese eingebetteten Dateien benützen die PDF-Datei als Container. Die eingebetteten Dateien können als *Attachment* oder als *Collection* eingebunden werden. Für die Darstellung von *Embedded Files* wird eine entsprechende Applikation benötigt z.B. für eine DOC-Datei ein Microsoft Word.

#### ***Attachment***

Das *Attachment* ist eine eingebettete Datei (*Embedded File*), welche vom Reader meistens als Büroklammer-Symbol dargestellt und darüber aufgerufen werden kann.

#### ***Collections***

Unter *Collections* wird die Funktion *PDF Portfolio* verstanden. Das *Portfolio* erlaubt eine Zusammenstellung verschiedener Dateien z.B. PDF-, DOC- und JPEG-Dateien zu einem "Dossier", wobei die Struktur und deren Komplexität frei wählbar sind.