

# Ingest-Level

Version 1.0

Datum: 16. 11. 2010

## INGEST-LEVEL

Dieses Dokument enthält sowohl eine Definition als auch eine Beschreibung der Vergaberegeln von Ingest-Leveln für digitale Objekte. Die Einführung einer Sortierung digitaler Objekte in Kategorien soll zu einer größeren Transparenz der angebotenen Dienstleistungen zwischen Dienstleister und Dienstnehmer beitragen.

### DEFINITION: INGEST-LEVEL

Ein Ingest-Level repräsentiert das Ergebnis eines mehrstufigen Prüfverfahrens, welches zurzeit durch vier aufeinander aufbauende Dienste innerhalb von DP4lib realisiert wird. Ein Ingest-Level wird jedem digitalen Objekt zugewiesen.

Der Ingest-Level ist eine Abschätzung bzgl. der wahrscheinlichen Langzeitarchivierbarkeit digitaler Objekte, die auf Grundlage momentan zur Verfügung stehender technischer Mittel und den etablierten Geschäftsprozessen des Dienstleisters getroffen wird. Der Ingest-Level stellt jedoch keine Aussage über die allgemeine Qualität der geprüften digitalen Objekte dar.

## KRITERIEN

### DATENINTEGRITÄT (DI)

Der vom Dienstnehmer übermittelte Bitstream des digitalen Objekts hat sich im Zuge der Datenübertragung nicht verändert.

### BESCHRÄNKUNGSFREIHEIT (BF)

Das digitale Objekt ist beschränkungsfrei, d. h. es existieren keine - seitens des Dienstleisters - technisch erkennbare Beschränkungen, welche die Nutzung des digitalen Objektes beeinträchtigen oder unmöglich machen.

### EXTRAKTION FORMATSPEZIFISCHER TECHNISCHER METADATEN (MD)

Formatspezifische technische Metadaten konnten generiert werden.

### FORMAT-VALIDITÄT (V)

Das Dateiformat des digitalen Objekts ist gegenüber einer Formatspezifikation valide.

## INGEST-LEVEL – EIGENSCHAFT – MATRIX

	DI	BF	MD	V
Level 0	X	O	O	O
Level 1	X	X	O	O
Level 2	X	X	X	O
Level 3	X	X	X	X

## BESCHREIBUNG DER INGEST-LEVEL – EIGENSCHAFTEN – MATRIX

Die Ingest-Levels in der obigen Tabelle bauen aufeinander auf. Je höher der Ingest-Level, desto größer wird – so die Hypothese – die Wahrscheinlichkeit, das digitale Objekt auf lange Sicht nutzbar zu halten.

### LEVEL 0

Das digitale Objekt erfüllt das Kriterium: Dateiintegrität

#### Beispiel:

Ein digitales Objekt erhält Level 0, wenn die Integrität des Bitstream nach der Übertragung an den Dienstleister erfolgreich, durch gegenseitig abgesprochene Prozesse, überprüft, bestätigt und protokolliert werden konnte.

### LEVEL 1

Das digitale Objekt erfüllt die Kriterien: Dateiintegrität und Beschränkungsfreiheit

#### Beispiel:

Ein PDF-Dokument wurde mithilfe des abgesprochenen Prüfverfahrens erfolgreich auf Dateiintegrität geprüft und ist von einem Analyse-Tool als beschränkungsfrei erkannt worden.

### LEVEL 2

Das digitale Objekt erfüllt die Kriterien: Dateiintegrität, Beschränkungsfreiheit und Extraktion formatspezifischer technischer Metadaten.

#### Beispiel:

Ein PDF-Dokument wurde mithilfe des abgesprochenen Checksummenverfahrens auf Dateiintegrität geprüft und ist von einem Analyse-Tool als beschränkungsfrei erkannt worden. Zudem konnten durch einen Metadatengenerator formatcharakteristische technische Metadaten erzeugt werden.

<b>LEVEL 3</b>
Des digitale Objekt erfüllt die Kriterien: Dateiintegrität, Beschränkungsfreiheit, Extraktion formatspezifischer technischer Metadaten und Format-Validität.
<p><b>Beispiel:</b>          Ein PDF-Dokument wurde mithilfe des abgesprochenen Checksummenverfahrens auf Dateiintegrität geprüft und ist von einem Analyse-Tool als beschränkungsfrei erkannt worden. Zudem konnten durch einen Metadatengenerator formatcharakteristische technische Metadaten erzeugt werden. Die Validität des Dateiformats konnte gegen eine PDF-Spezifikation bestätigt werden.</p>

**ANMERKUNG:**

Aus Gesprächen mit den Partnern wurde deutlich, dass es evtl. auch zu überlegen ist, ob nicht noch ein weiteres Ingest-Level zu vergeben ist. Dieser würde einen expliziten Identifikationsschritt erforderlich machen, der ausschließlich die Aufgabe hat das Dateiformat zu identifizieren.

**IDENTIFIKATION (ID)**

Das digitale Objekt wurde hinsichtlich seines Dateiformats identifiziert.

Die angepasste Ingest-Level – Eigenschaft – Matrix würde dann wie folgt aussehen:

	DI	ID	BF	MD	V
<b>Level 0</b>	X	O	O	O	O
<b>Level 1</b>	X	X	O	O	O
<b>Level 2</b>	X	X	X	O	O
<b>Level 3</b>	X	X	X	X	O
<b>Level 4</b>	X	X	X	X	X

Über diese Tabelle und die Ingest-Levels würden wir gerne mit Ihnen am Donnerstag unter „TOP Muster-Kategorien“ reden. Anbei finden Sie eine Muster-Kategorisierung in die Ingest-Levels, dessen Einteilung nach Level 0-3 erfolgt.

## MUSTER-KATEGORISIERUNG IN INGEST-LEVEL

Diese Kategorisierung zeigt in der ersten Spalte die momentan relevanten Dateiformate. Die zweite Spalte gibt den maximal möglichen Ingest-Level wieder. Da die Ingest-Level aufeinander aufbauen wird nur der höchstmögliche Ingest-Level notiert. In der letzten Spalte können Sie den von Ihnen angestrebten Ingest-Level eintragen.

	Möglicher Ingest-Level	Wunsch
<b>PDF</b>		
Version 1.0	Level 3	
Version 1.1	Level 3	
Version 1.2	Level 3	
Version 1.3	Level 3	
Version 1.4	Level 3	
Version 1.5	Level 3	
Version 1.6	Level 3	
Version 1.7	Level 0	
Tagged	Level 3	
Linearized	Level 3	
PDF/A-1 a	Level 3	
PDF/A-1 b	Level 3	
<b>TIFF</b>		
Version 4.0 – 6.0	Level 3	
TIFF/EP	Level 3	
Baseline 6.0 Class B, G, P and R	Level 3	
<b>GIF</b>		
87a	Level 3	
89a	Level 3	
<b>JPEG</b>		
ISO/IEC 10918-1:1994	Level 3	
JFIF 1.2	Level 3	
Exif 2.0 - 2.2	Level 3	
SPIFF	Level 3	
JTIP	Level 3	
JPEG-LS	Level 3	

<b>ASCII</b>		
ANSI X3.4-1986	Level 3	
ECMA-6	Level 3	
ISO649:1991	Level 3	
<b>XML</b>		
1.0	Level 3	
<b>Microsoft Office</b>		
doc,xls,ppt	Level 0	
<b>HTML</b>		
3.2	Level 3	
4.0	Level 3	
4.01	Level 3	
<b>ZIP</b>		
	Level 0	
<b>AIFF</b>		
1.3	Level 3	
C	Level 3	
<b>WAVE</b>		
PCMWAVE	Level 3	
FORMATEX	Level 3	
FORMAT EXTENSION	Level 3	
<b>CDImages</b>		
ISO-9660	Level 2	
<b>PS</b>		
	Level 2	