

Gregor Fellenz

Mediengestalter für Digital- und Printmedien
Studium Druck- und Medientechnologie, HdM Stuttgart
pagina GmbH, Tübingen

Lehrauftrag InDesign Satzautomation, HdM Stuttgart
www.indesignblog.com

gregor.fellenz@netzfrei.org

Übersicht

1. Technologie SnapShot XML

2. InDesign und XML

- Wozu und für wen eigentlich?
- Voraussetzungen.
- Arbeiten mit Bordmitteln.
- Einschränkungen und Probleme.

3. Automatisierung

- Erweiterte Möglichkeiten.
- XML-Rules kurz vorgestellt.

4. Datenbanken

- Datenbankbindung vs. XML.

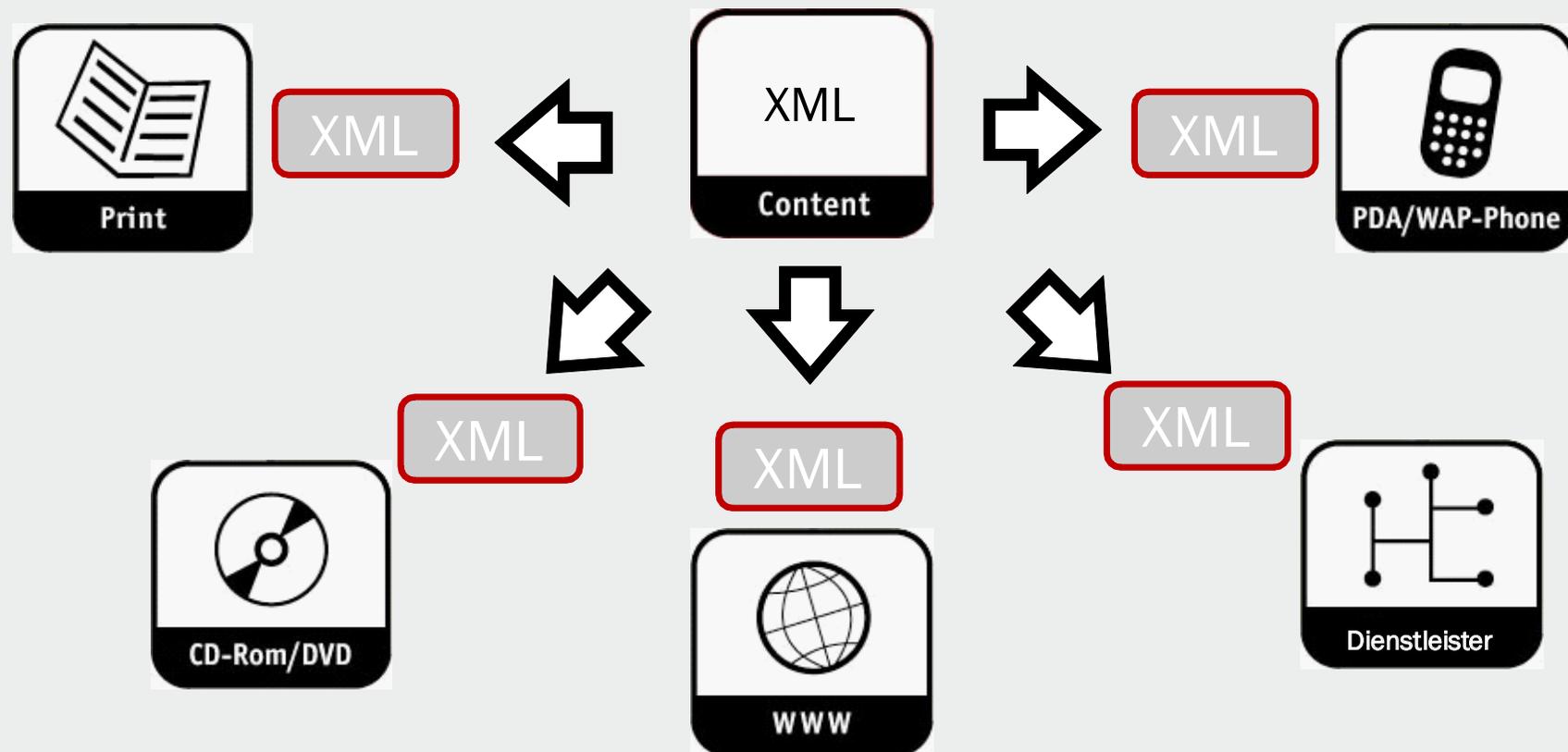
Technologie SnapShot XML

Warum überhaupt XML

- **Datenchaos**
Formate, Datenprobleme, Konvertierungsprobleme
- Keine **Wiederverwertbarkeit**
Langzeitarchivierbarkeit, Kompatibilität
- **Mehrfachverwertung** der Daten
Satz, Internet, DVD Nachschlagewerke, Mobile Applications
- **Automatisierte Produktion**
- **Standardisierung im Satz**

Aus einer Quelle in alle Zielmedien

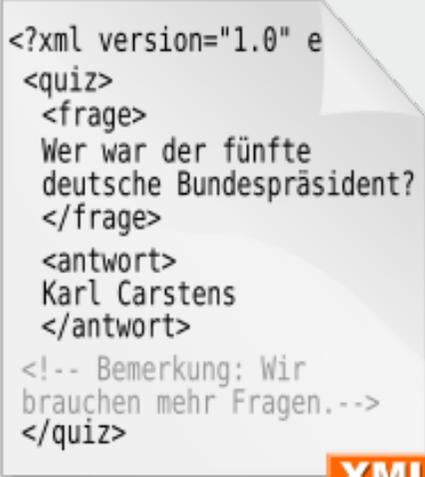
- Herausforderung: Inhalte aus dem Layoutsystem nach XML zurück überführen (**Roundtripping**).
- Alternative: Medienspezifisch bearbeiten und speichern (**Doppelpflege**).



Extensible Markup Language

Standardisiertes Dateiformat

- Austausch von Dokumenten, Definition von Datenstrukturen
- Sprache zur Beschreibung von **Dokumentstrukturen** bzw. eine Metasprache, die eine abstrakte Syntax für Auszeichnungssprachen definiert
- **Dokumentzentriert, datenzentriert**
- **Keine Datenbank**



```
<?xml version="1.0" e
<quiz>
  <frage>
    Wer war der fünfte
    deutsche Bundespräsident?
  </frage>
  <antwort>
    Karl Carstens
  </antwort>
  <!-- Bemerkung: Wir
  brauchen mehr Fragen.-->
</quiz>
```

XML

Auch im Bereich Publishing?

Mit **Strukturierung** wurde schon immer das **Erfassen des Informationsgehaltes** erleichtert.

Umgesetzt durch **Formatierung** wird indirekt die Struktur vom Leser wahrgenommen.

Aber: XML ist nur eine Sprache

Die Kunst, Information zu strukturieren, wird durch XML nicht ersetzt.

Fachbegriffe

▪ **Element**

Im Dokument mittels

- passendem Paar aus **Start-Tag** (<Tag-Name>) und **End-Tag** (</Tag-Name>) oder
- mit einem **Empty-Element**-Tag (<Tag-Name />) ausgezeichnet.

▪ **Attribut**

Für Zusatz-Informationen zu Elementen (parameter="wert"), Meta-Informationen.

- **Verarbeitungsanweisungen** (<?Ziel-Name Parameter ?>, engl. Processing Instruction)
- **Kommentare** (<!-- Kommentar-Text -->)
- **Namensräume** (<html xmlns="http://www.w3.org/1999/xhtml" xmlns:svg="http://www.w3.org/2000/svg">) Elemente unterscheiden.

Ein XML-Dokument muss genau ein Element (**Root**) auf der obersten Ebene enthalten. Unterhalb von diesem Dokumentelement können weitere Elemente **verschachtelt** werden.

- Eine **Document Type Definition** (DTD) ist ein Regelwerk für XML-Dateien. Reihenfolge der Elemente und welche überhaupt vorkommen dürfen.
- XML-Dokumente sind idealerweise in **utf-8** codiert.

InDesign und XML

*InDesign ist eine von **vielen Applikationen** die XML unterstützt.*

- **Richtig erstellte** InDesign-Dokumente können als XML weiterverwendet werden.
- Nach **entsprechender Anpassung** kann InDesign beliebige XML-Dokumente formatieren.

Wozu und für wen eigentlich?

- Dokumente **liegen als XML** vor.
- Ähnliche Strukturen in **wiederkehrenden** Aufträgen/Projekten.
- Dokumente sollen **weiterverwendet** werden.
- Daten müssen **als XML abgeliefert** werden.
- Bereitschaft sich mit der **Technologie** auseinanderzusetzen.



Voraussetzungen

- XML muss **wohlgeformt, bzw. valide** sein.
- Struktur muss in InDesign **abbildbar** sein.
 - Attribute vs. Elemente.
 - Strukturen/Kapitel nur eingeschränkt abbildbar.
 - Whitespace-Handling.

Strategische Überlegungen

- Know-how im Unternehmen/Dienstleister vorhanden?
- Grad der **Automatisierung**.
- Anforderungen an den XML-Export.
 - Single-Source-Publishing
 - Roundtripping



Typischer Workflow



- **Wo** kommen die Daten her?
- Ist **Roundtripping** gewünscht?
- Welche **Zielformate** sollen bedient werden?

Oder nur Teilstrecken...

- XML-Dokumente vorhanden → **Import**
- InDesign Dokumente vorhanden → **Export**

Arbeiten mit Bordmitteln

Einfache Möglichkeiten zum Erstellen und Bearbeiten von XML-Dokumenten sind in InDesign vorhanden.

InDesign Dokumente richtig bearbeiten

- **Absatzformate**

Mit Absatzformaten können Layoutoptionen auf Absätze angewendet werden.

- **Zeichenformate**

Textstellen oder einzelne Zeichen können durch Zeichenformate modifiziert werden.

- **Tabellenformate**

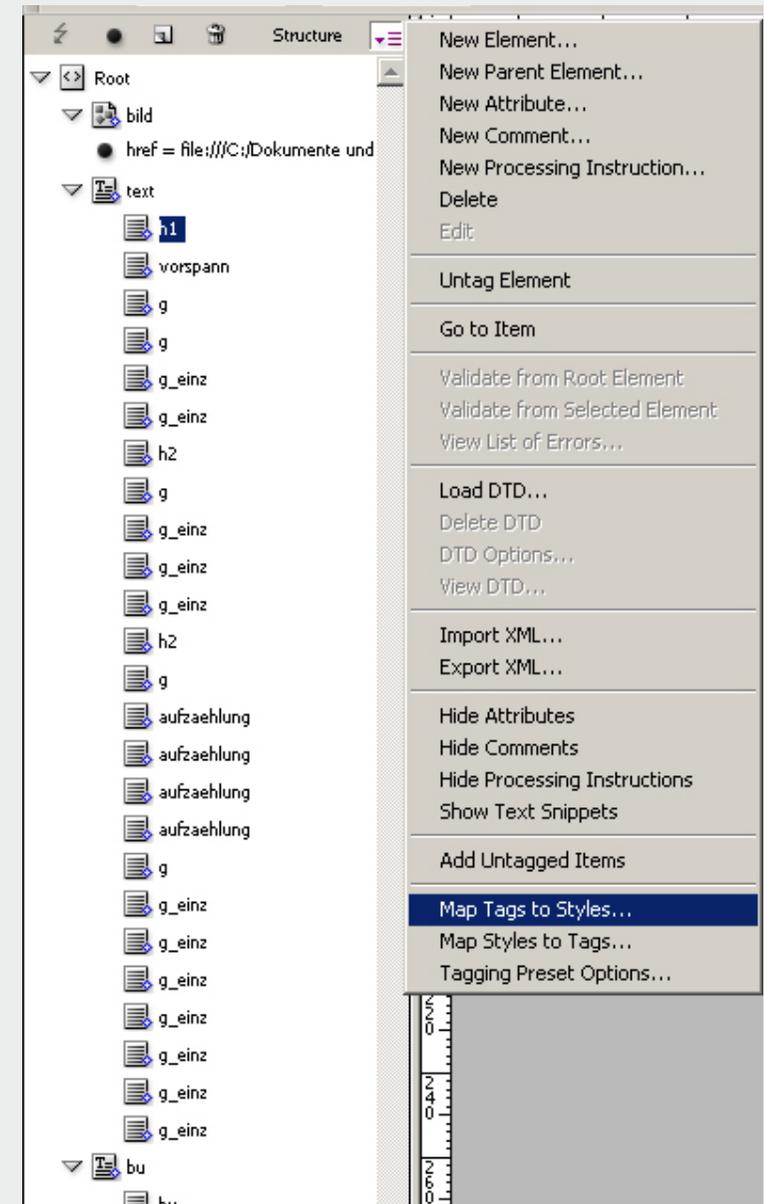
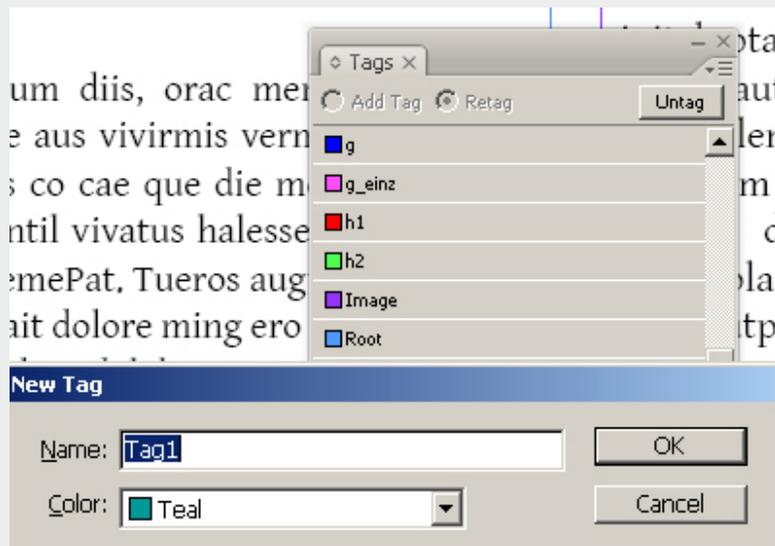
- **Fußnoten und Verweise**

InDesign Funktionen nutzen (Datenqualität).

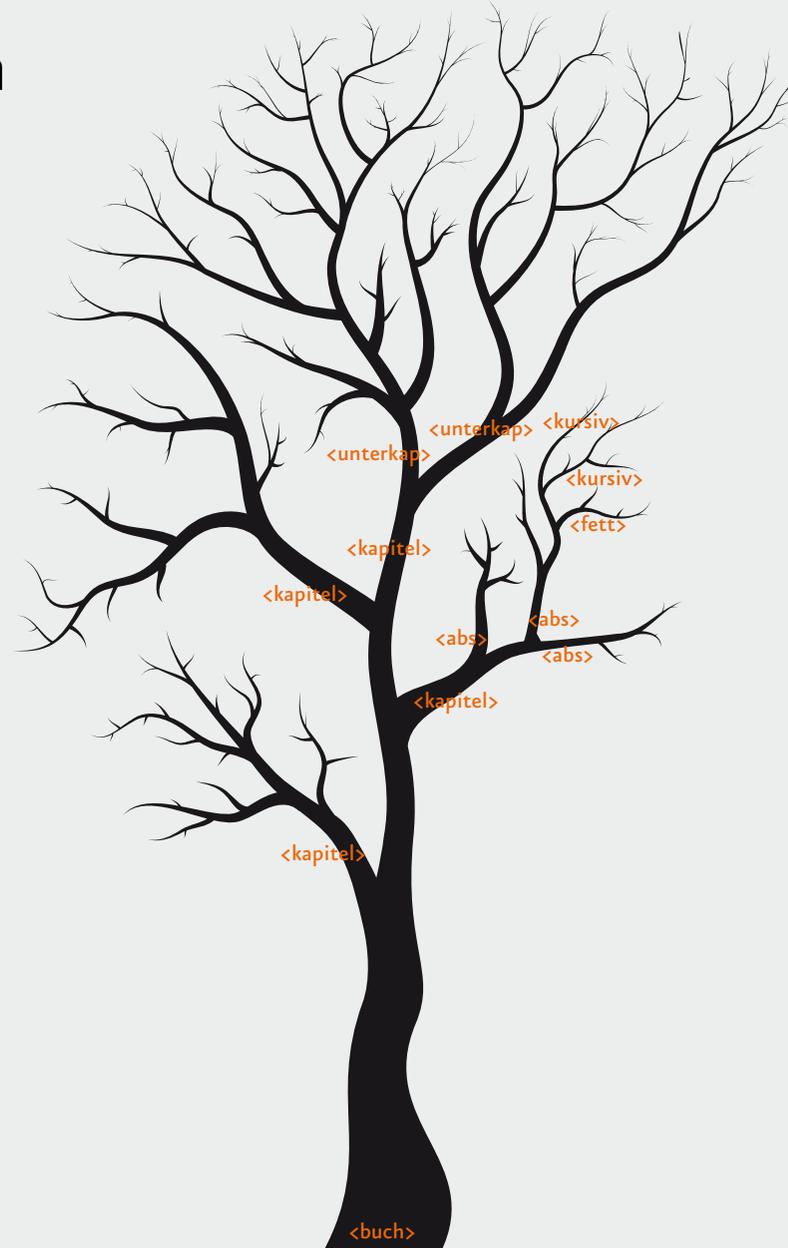
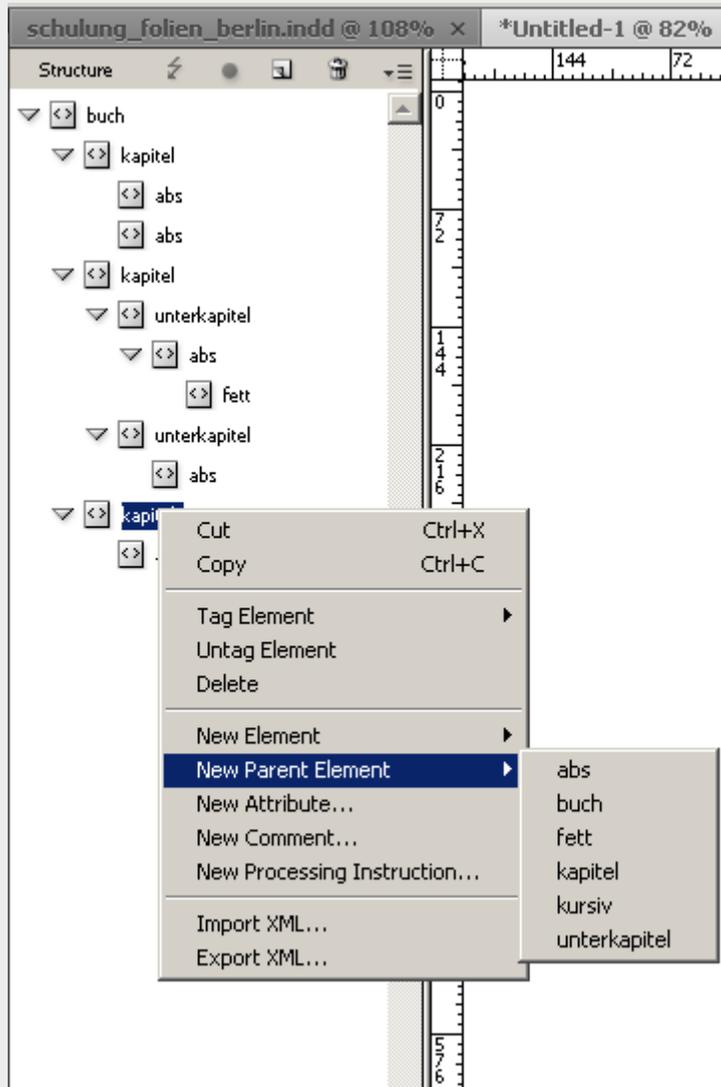


XML-Ansichtsoptionen

- Strukturansicht
- Tag-Marken/
Rahmen mit Tags
- Textansicht
- **Tags**
 - **Elementnamen** können auch direkt erstellt und zugewiesen werden.
 - **Attribute** sind nicht so intuitiv.



Strukturansicht und XML-Baum



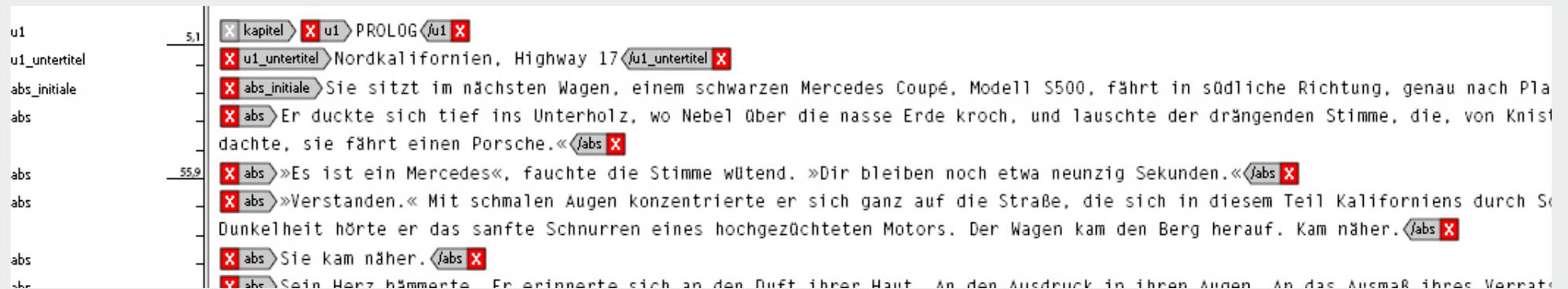
Blocklevel Elemente/Absätze

- Blocklevel Elemente werden als **Absatzformatvorlagen** umgesetzt
- Wenn Elemente in den XML-Daten erhalten bleiben müssen:
Doppelpflege!



XML bearbeiten

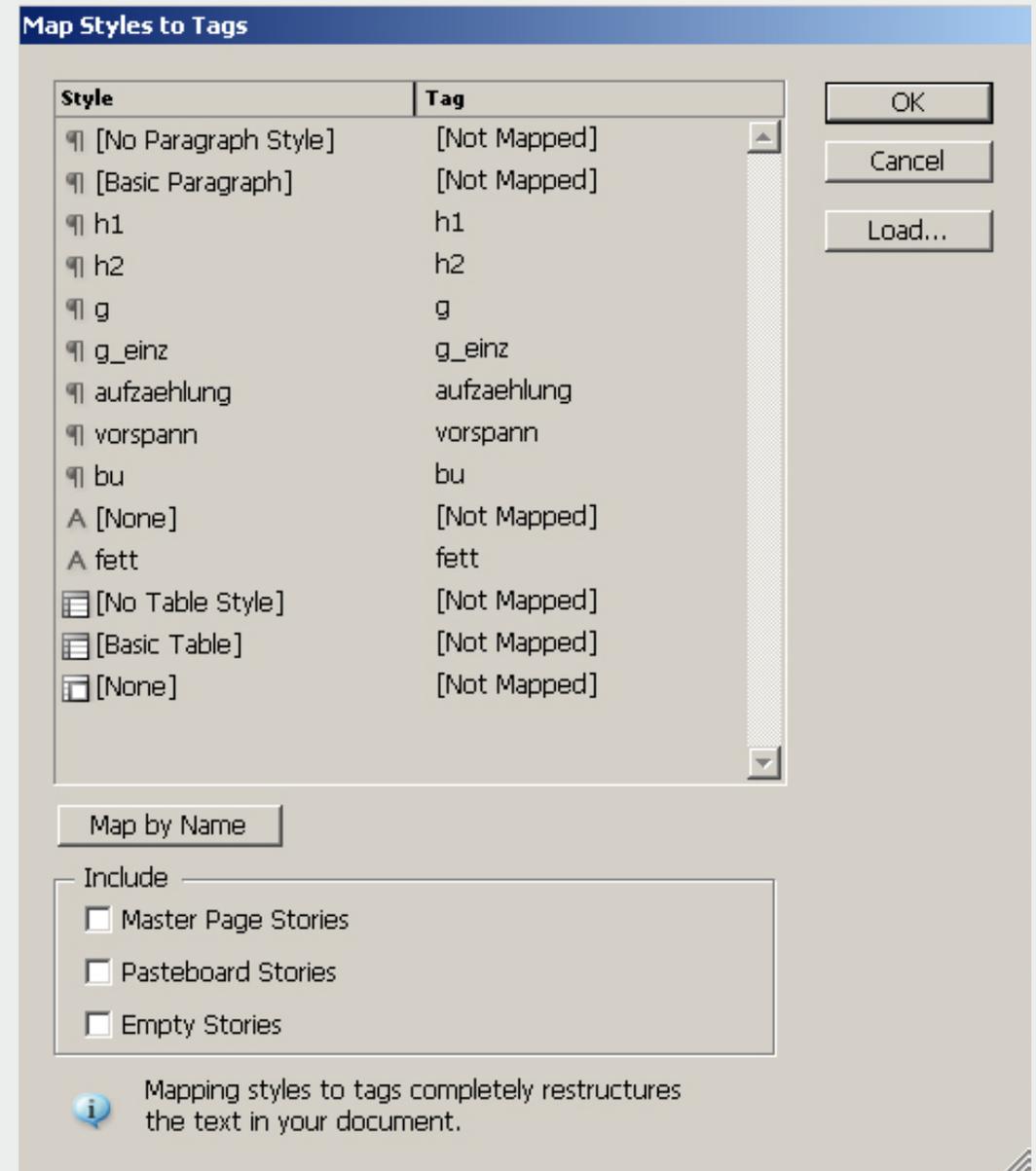
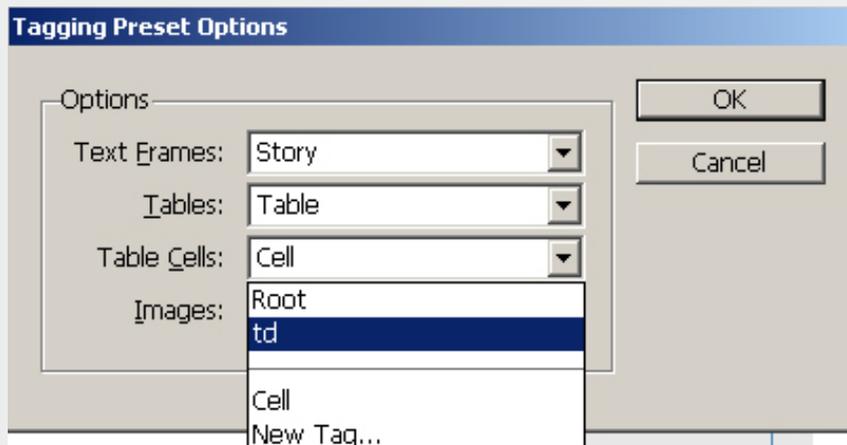
- Am besten im **Texteditor** mit eingeblendeten Tag-Marken



- In der Tags-Palette muss der Radio Button **Neu zuordnen** angewählt sein
- Cursor muss in einem Element stehen
Zwischen Elementen wird das Eltern-Element umgetaggt (z.B. <kapitel> -> <abs>)

Auszeichnung

- **Formate zu Tags** zuordnen
(oder vice versa)
Formaten werden Tags zugeordnet.
- **Zuweisungen** für Tabellen, Textrahmen
und (Bilder)

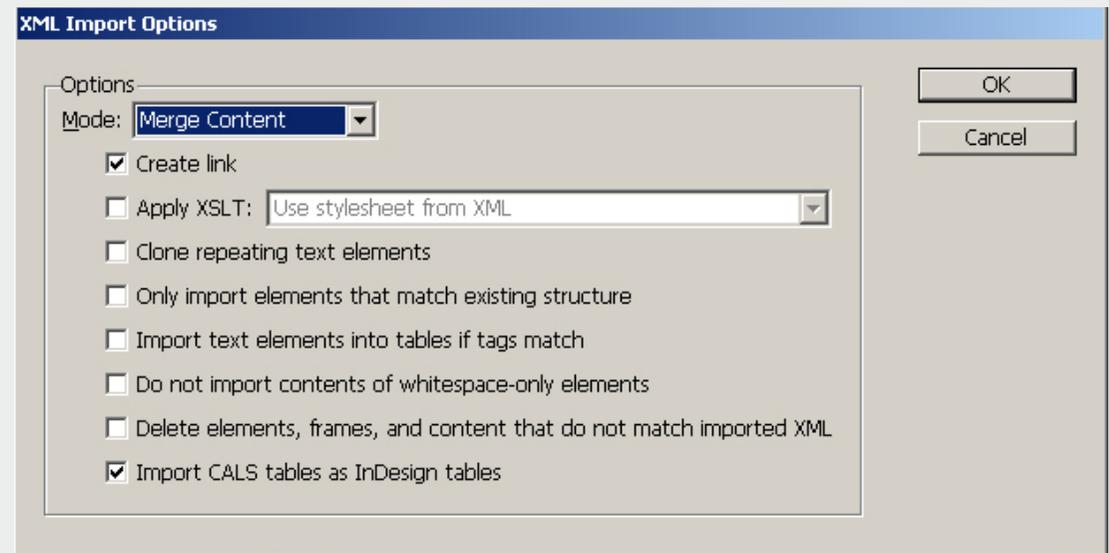


Praxisbeispiel

- Erstellen/Bearbeiten von XML Dateien in InDesign
 - Aufbau eines Dokuments
 - Struktur- und Textansicht
 - Zuweisung von Tags
 - Bearbeiten von Inhalten
- Export von XML (Bilder, Tabellen?)
 - Exportoptionen
 - Spezialzeichen gehen verloren

XML Import

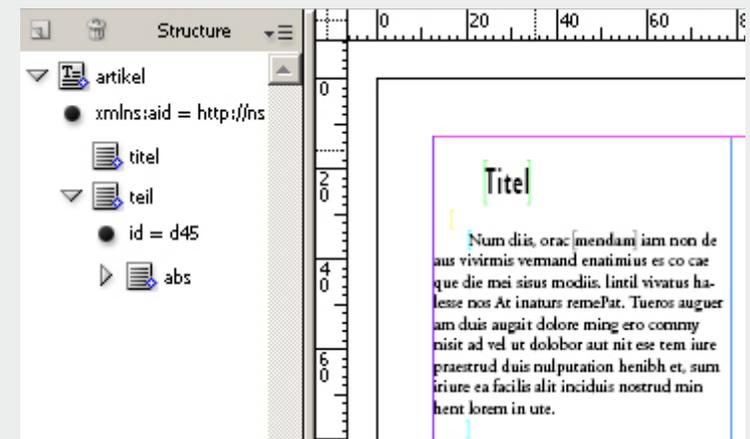
- **Einfache Strukturen** werden durch Formatzuweisungen gestaltet.
- Geschickte Lösungen möglich aber wenig flexibel.
- **Vorteil:** Ohne Programmierung oder Plugins.



Whitespace Handling

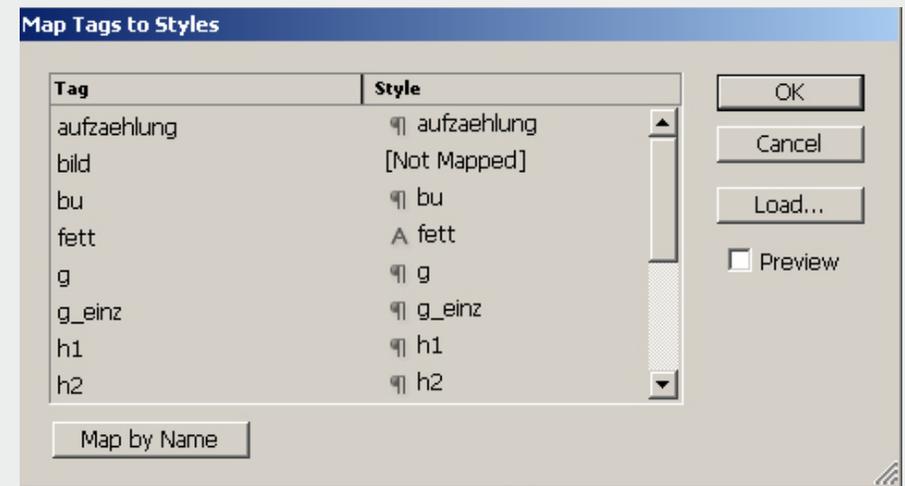
- **Whitespace** wird vollständig **importiert** und dargestellt!
- **Oder** wird vollständig ignoriert.
- **Lösung** Importskript verwenden.
Download auch unter: <http://www.indesignblog.com/>

```
<?xml version="1.0" encoding="UTF-8"?>
<artikel>
  <titel>Titel</titel>
  <teil id="d45">
    <abs>Num diis, orac <fett>mendam</fett>
```



Formatierung

- **Tags zu Formaten** zuordnen (oder vice versa)
Einzelne Elemente können Formaten zugeordnet werden.
- **Bilder** und **Tabellen** automatisch formatieren.
- Elementen über **Attribute** Styles zuordnen.
- XSL-Transformationen und XML-Rules.



```
<artikel xmlns:aid="http://ns.adobe.com/AdobeInDesign/4.0/" >
  <titel aid:pstyle="h1">Titel</titel>
  <teil id="d45">
    <abs aid:pstyle="g">Num diis, orac <fett aid:cstyle="fett">mendam</fett> iam
    non de aus vivirmis vermand enatimius es co cae que die mei sisus modiis. lin-
    til vivatus halesses nos At inaturs remePat. Tueros auguer am duis augait dolo-
    re ming ero commy nisit ad vel ut dolobor aut nit ese tem iure praestrud duis
    nulputation henibh et, sum iriure ea facilis alit inciduis nostrud min hent
    lorem in ute.</abs>
  </teil>
</artikel>
```

Praxisbeispiel

- Import von XML
 - Standardfunktionalität
 - Import mit Zusatz-Skript
 - Formate zu Tags zuordnen

Tabellen

InDesign unterstützt das CALS-Tabellenformat und sein eigenes Tabellenformat.

- **InDesign Tabellenformat**

Eher Softwarenah? In anderen Kontexten aufwändig Konvertierung notwendig.

- **CALS Tabellen** (nur CS3)

OASIS Standard, kein xhtml aber Konvertierung vorhanden.

www.oasis-open.org/specs/tablemodels.php

- Formatierung über Attribute oder Formatzuweisungen (nur CS3).

- Struktur von InDesign Tabellen.

```
<Table xmlns:aid="http://ns.adobe.com/AdobeInDesign/4.0/" aid:table="table"
aid:trows="6" aid:tcols="2">
  <Cell aid:table="cell" aid:thead="1" aid:crows="1" aid:ccols="1"
aid:ccolwidth="67"><t_header>Verwendung</t_header></Cell>
  <Cell aid:table="cell" aid:thead="1" aid:crows="1" aid:ccols="1"
aid:ccolwidth="87"><t_header>Bauart</t_header></Cell>
  <Cell aid:table="cell" aid:crows="1" aid:ccols="1" aid:ccolwidth="67">
<t_abs>Industriesauger</t_abs></Cell>
  ...
</Table>
```

Entity Auflösung

Alle Entities werden beim Import aufgelöst.

- **Standard Entities**
& = & | < = < | > = >
- Character Entities für **UTF-8 Zeichen**
&#x####;
<http://www.decodeunicode.org/>
- Für eigene Entities muss eine **DTD verknüpft** sein
oder die Entity muss im **Dokument deklariert** sein.
<!ENTITY idcs „InDesign CS3“>
Aufruf im Dokument über &idcs;
- Bei nicht auflösbaren Entities **bricht der Import** ab!



< CHARACTER > BLOCK PROPERTIES
U+20B0 GERMAN PENNY SIGN

Bilder

Bilder werden automatisch gesucht wenn ein href Attribut vorhanden ist

- **Rudimentäres Handling** über href Attribut.
 - Nur Import
 - Keine Größen, Ausschnitte etc.
- Nur **verlinkte Bilder** möglich – Einbetten ist grundsätzlich nicht empfehlenswert.

```
<bild href="file://path/images/testfile.jpg" />
```

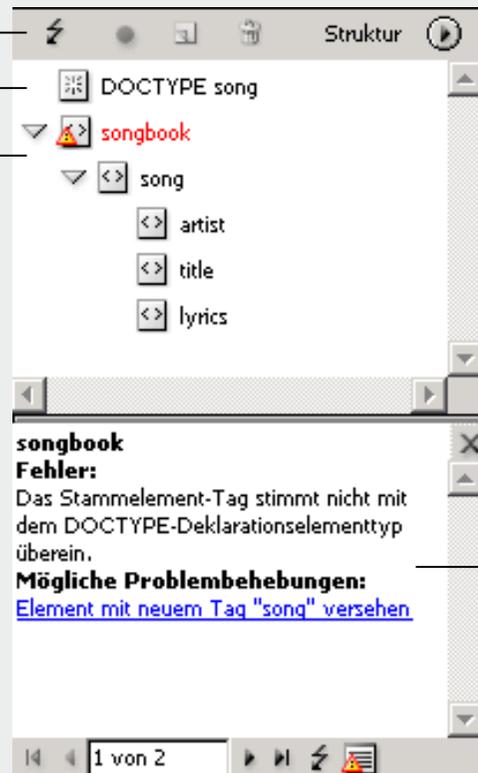
Weitere XML-Funktionen

- DTD importieren
InDesign kann eine DTD (Document Type Definition) importieren und gegen diese validieren.
- DTD validieren
Ab dem Stammelement (Root-Element)
oder ab einem ausgewählten Element gegen die DTD validieren.

Struktur validieren

Name der DTD

Struktur des
Dokuments



Fehlerliste

XML Export aus InDesign

- **Remap Special Characters**

Sonderzeichen werden als Leerzeichen ins XML geschrieben.

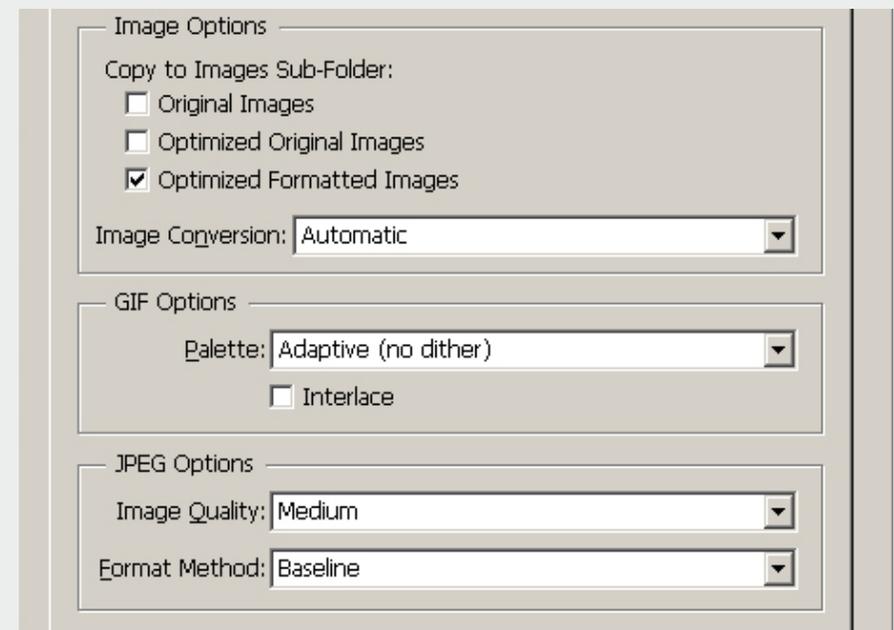
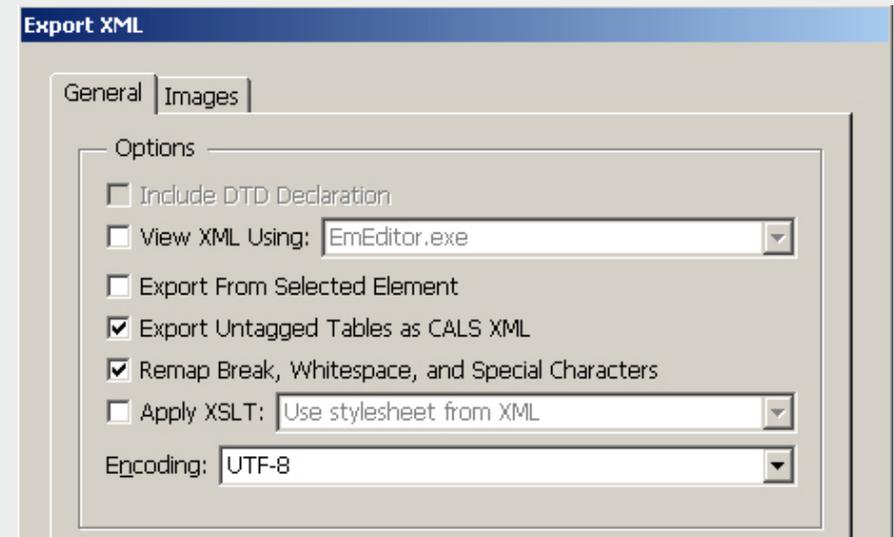


- **Tabellen**

Tabellen können als CALS-Tabellen exportiert werden

- **Bilder**

Werden auch exportiert, Downsampling möglich.



XHTML-Export

- Skript von Adobe zum Export von InDesign Dokument als **XHTML**.
- Einfache Möglichkeit um Inhalte zu generisch zu **extrahieren**.
- Absatz- und Zeichenformate werden als CSS class Attribute übergeben.
- **Tabellen**, Fußnoten und Verweise werden exportiert!
- Listen aus Absatzformaten werden konvertiert.

Digital Editions/epub-Export

- Basiert auf XHTML Export
- **Nicht für komplexe Layouts geeignet**
 - ohne Nachbearbeitung können nur einfache Dokumente exportiert werden.
- Verbesserungen in CS4

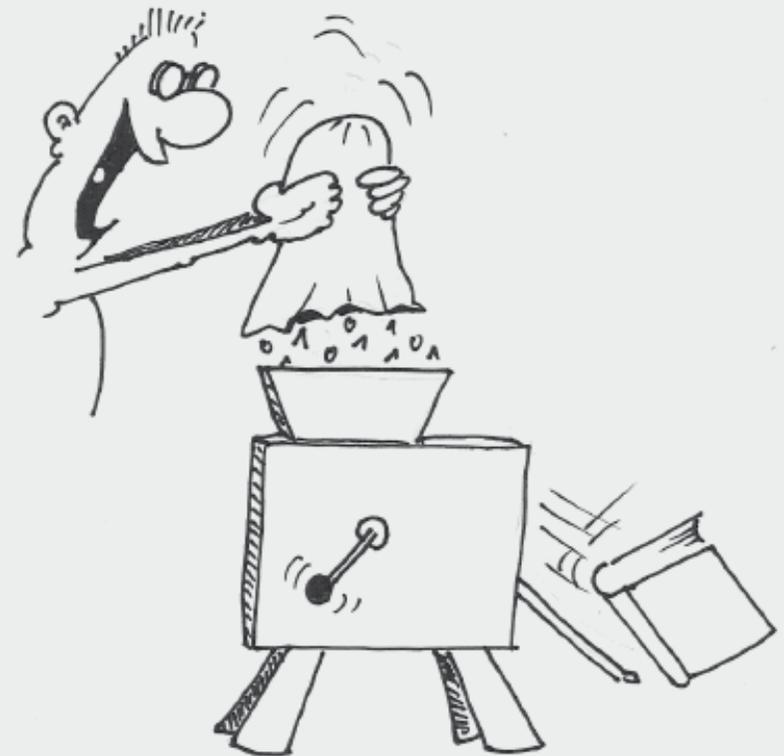
Einschränkungen und Probleme

- Struktur muss in InDesign **abbildbar** sein.
- XML-Daten müssen **vorhanden** sein, bzw. **generiert** werden können.
- Wirtschaftlicher Nutzen?
- **Tabellenunterstützung bis CS3** suboptimal.
- Hierarchie-Ebenen mühsam.
- **Bearbeitungsmöglichkeiten** verbesserungswürdig.
- **Export** nicht optimal
Nachbearbeitung oftmals notwendig.
- Geschwindigkeit.

Automatisierung

Erweiterte Möglichkeiten

- Keine Grenzen mehr fürs Layout – Aufbau **komplexer Strukturen** steuerbar.
- Satz von wiederkehrende Aufgaben entlasten – mehr Zeit fürs Layout.
- Verweise und Indexerstellung möglich.
- **Scripting** oder Plugin-Entwicklung.
- Scripting in VB Skript, AppleScript und JavaScript.
- Adobe Scripting Guide.



XML-Rules kurz vorgestellt

XML-Rules eignen sich dazu dynamisch Layouts zu erstellen und formatieren.

- **XML-Rule-Sets** sind kleine Skripte die bei bestimmten Knoten Aktionen auslösen.
- **XPath** gesteuert.
- Wenn die Regel zutrifft, wird eine bestimmte Aktion ausgelöst.
- VB Skript (Windows), AppleScript (MacOS) oder JavaScript (beide).
- **Beispiele**
 - Eine Regel für Absatzelemente überprüft, ob es der erste Absatz nach einer Überschrift ist und rückt ihn entsprechend ein.
 - Bei einer bestimmten Hierarchie-Ebene wird eine neue rechte Seite erstellt.

```
function formatTelefonnummer(){
  this.name = „formatTelefonnummer“;
  this.xpath = „//TelNummer“;
  this.apply = function(myElement, myRuleProcessor){
    with(myElement){
      insertTextAsContent(„Telefon:\t“, XMLElementPosition.ELEMENT_START);
      applyParagraphStyle(_dokument.paragraphStyles.item(„TELNUMMER“));
    }
  }
  return true;}
}
```

Datenbanken

Datenbankanbindung vs. XML

- **Datenbanklösungen** oft für Katalogproduktionen.
- Publishing Lösung wird in die Datenbank integriert.
- XML flexibler, für „kleine“ Workflows ausreichend.
- Der kleine Bruder: Data Merge (CSV, TXT).
- Datenbank → XML → InDesign.

Plugins

- Für XML nicht zwingend nötig.
- **Datenbank Plugins** sollten sehr gut automatisierbar sein.
- Relativ teuer.
- **Übersicht** von Jörg Oyen:
http://publishing-report.com/database_publishing_tools.html

Literatur, Informationen und Quellen

InDesign und XML

- Adobe Informationen zu XML
http://www.adobe.com/designcenter/indesign/articles/indcs3ip_xmlrules.pdf

InDesign Scripting

- Adobe Scripting Guide
<http://www.adobe.com/devnet/indesign/>
- Adobe Scripting Forum (sehr rege Community)
http://forums.adobe.com/community/indesign/indesign_scripting

Datenbank Plugins

- **Übersicht** von Jörg Oyen:
<http://www.publishing-report.com/marktuebersichten/>

Vielen Dank für Ihre Aufmerksamkeit!

Fragen? Anregungen? Diskussionsbedarf?

Die Folien zum Download:

<http://www.indesignblog.com/tekom>

E-Mail:

gregor.fellenz@netzfrei.org