

# Whitepaper parsX

## Konzept und Module

Stand: Februar 2018



*Kann es den einen, optimalen Verlagsworkflow geben – bei so unterschiedlichen Verlagsprogrammen, Unternehmensgrößen und Geschäftsmodellen? Während der eine Verlag gerne an den bewährten Abläufen festhalten und die Daten- und E-Book-Produktion gerne in den Händen seiner Dienstleister wissen möchte, plant der andere Verlag, die Produktionsprozesse im eigenen Haus zu optimieren und über ein Workflowmanagement- bzw. Contentmanagement-System zentral zu steuern.*

*Mit parsX hat pagina ein Set von Softwaremodulen entwickelt, um sämtliche Arbeitsschritte von der Manuskriptphase bis zur Ausgabe in alle Medienkanäle komfortabel und qualitätsgesichert vornehmen zu können – gleichgültig, ob dies im Verlag selbst oder beim Dienstleister geschieht. Sämtliche Prozesse sind XML-basiert und fußen auf einer Datenstruktur, die allen Ansprüchen einer zukunftsgerichteten Content-Strategie genügt. Neben der Erstellung optimaler Verlagsprodukte ist der Fokus der Weiterentwicklung von parsX vor allem auf innovative Vertriebskonzepte und die Unterstützung digitaler Verlagsstrategien gerichtet. parsX läuft gleichermaßen im Filesystem (also ohne CMS oder MAM-System) wie als Produktionsmodul innerhalb datenbankgestützter Publikationssysteme.*

*Zentraler Ansatz bei parsX ist ein moderner Kooperationsgedanke sowie das Schaffen einer starken technischen Gemeinschaft von Publikumsverlagen, um in den für Verlage wirtschaftlich schwierigen Zeiten eine starke Lobby zu bilden, voneinander zu lernen und gemeinsam neue Technologien entwickeln zu können. Nachfolgend werden diese Ansätze beschrieben.*

## Die Maximen bei der Entwicklung von parsX

### 1. Unsere Mission: Die Etablierung eines Branchenstandards für Publikumsverlage

In Industrie und Wissenschaft hat sich längst die Erkenntnis durchgesetzt, dass der Einsatz eines branchenspezifischen Datenstandards der Schlüssel zum Erfolg in der digitalen Zeit ist. Ob Automobilbranche, Chemieindustrie, universitäre Forschung oder Börsenhandel: Erst die Etablierung eines Branchenstandards erlaubt schnellen Datenaustausch, die Entwicklung von Standardsoftware und den Eintritt in das „Industrie 4.0-Zeitalter“. Auch für zahlreiche Verlagssparten gibt es mittlerweile XML-Standards: *TEI* für die geisteswissenschaftlichen Verlage, die *NLM-DTD* für naturwissenschaftliche Publikationen, *DITA* für technische Dokumentationen und vieles mehr.

Nur für die Anforderungen der Publikumsverlage gab es bislang keine erprobte und bewährte XML-Struktur, obwohl die Anforderungen täglich steigen. Solange aber jeder Verlag seine eigene Struktur entwickelt, haben die Publikumsverlage keine Lobby; es gibt keine schnell implementierbaren Softwarelösungen von Drittanbietern, keine kostengünstige Implementierung in CMS- oder MAM-Systeme.

Mit parsX wollen wir das ändern.

parsX ist der Datenstandard für Publikumsverlage. Die zugrundeliegende DTD wurde in enger Abstimmung mit großen, innovativen Verlagshäusern entwickelt und beinhaltet die Architekturen für alle Anforderungen eines crossmedial ausgerichteten Publikumsverlags. Der modulare Aufbau stellt sicher, dass Sie dabei nur die Elemente angeboten bekommen, die zu Ihrem Verlagsprogramm passen. In der Regel sind keinerlei Anpassungen erforderlich, um Ihr Verlagsprogramm optimal abzubilden. Wenn Ihre Anforderungen steigen, erlaubt das Semantik-Modul jederzeit die Anpassung an verlagsspezifische Gegebenheiten, ohne die Basisstruktur zu gefährden.

Immer mehr Softwareentwickler und Dienstleister im In- und Ausland bieten Lösungen und Dienstleistungen auf der Grundlage von parsX-Daten an. Und immer mehr Anbieter von CMS/MAM-Systemen integrieren die parsX-Module in ihre Lösungen.

## 2. Unsere Überzeugung: Partnerschaft statt Konkurrenz Systematische Einbindung Ihrer Dienstleister

Wer die Medienbranche verändern will, darf keine Angst vor Mitbewerbern haben. Wir sind der Überzeugung: Eine Softwarelösung, die einen Verlag an einen Dienstleister bindet, ist unzeitgemäß und strategisch falsch. Deshalb erwerben Sie mit Ihrer parsX-Lizenz automatisch das Recht, die Software auch bei Ihren Vorstufen-Dienstleistern einzusetzen – ohne dass diesen (von einer geringen jährlichen Support-Gebühr abgesehen) dafür eigene Kosten entstehen. Einzige Bedingung ist der Abschluss eines Unterlizenzvertrages, der sicherstellt, dass der Dienstleister die Software nur für parsX-Lizenznehmer einsetzt. Ein Qualifizierungsprogramm stellt das XML-Knowhow und den souveränen Umgang mit parsX bei Ihren Dienstleistern sicher. So wollen wir dazu beitragen, dass der Wandel in der Branche tatsächlich gelingt und nicht am „Klein-klein“ im Konkurrenzverhalten zwischen Ihren Satzpartnern behindert wird. Schon ca. 30 Dienstleister arbeiten heute nach dem parsX-Standard, ständig kommen neue hinzu. Wahrscheinlich ist parsX schon bei einem Teil auch Ihrer Vorstufenpartner im Einsatz. Selbstverständlich steht auch pagina Ihnen als Konvertierungs- und Satzdienstleister mit langjähriger XML-Erfahrung zur Verfügung – aber eben nicht, weil Sie keine andere Wahl hätten. Sondern nur, wenn Sie mit uns zusammenarbeiten wollen.

## 3. Unser Ziel: Der Publikumsverlag 4.0

Immer wieder erlebt unsere Branche technologische Erneuerungen. Während sich aber alle bisherigen Technologiesprünge stets auf den Beschaffungsmarkt bezogen und damit eingekauft werden konnten, erlebt die Branche derzeit den ersten großen Wandel im Absatzmarkt – mit allen offenen Fragen.

Wir von pagina sind uns sicher: Das Verlagsgeschäft der Zukunft kann nicht darin bestehen, statt einem Medium – dem Buch – künftig eben zwei Medienformen – Buch und E-Book – in Auftrag zu geben. Vielmehr sind, ohne das Kerngeschäft aus den Augen zu verlieren, innovative Vertriebsmodelle, flexibles Agieren im Autorenmarkt, E-Book-first-Konzepte und vieles mehr gefragt. Mit parsX wollen wir Antworten geben. Mit Modulen, die nicht nur die optimierte Satz- und E-Book-Produktion betreffen, sondern darüber hinaus eine automatisierte Leseproben-Produktion ermöglichen, Multimediale E-Books erstellen, die Seitenwechsel aus InDesign ausspielen (zur Erzeugung von EPUB-PageLists) oder auch beliebige Leseproben zu einem E-Book für den Vertrieb aggregieren. Automatisches Splitten von Romanen in einzelne Serials, ein Qualitätssicherungskonzept, das keine eigenen XML-Kenntnisse erfordert, eigene Copy-Editing-

Module in Word für einen E-Book-first-Workflow, individuell hinterlegbare E-Book-Profile, vollautomatische Satzvorschau und vieles mehr: Alle parsX-Module werden mit dem Ziel entwickelt, konkrete Lösungen für Aufgabenstellungen im Publikumsverlag anzubieten. Und alles zum Einsatz sowohl im Verlag als auch bei Ihren Dienstleistern. Dabei soll die Umstellung eines Verlags auf eine XML-basierte Produktion – inklusive aller Schulungen – innerhalb weniger Monate möglich sein.

Im Rahmen der Maintenance erhalten Sie regelmäßig neue Releases, z.T. auch mit neuen, für sie kostenlosen Modulen. Und wenn Sie selbst Anregungen für ein neues Modul oder eine Erweiterung haben, sprechen Sie uns an. Wir prüfen und kalkulieren jede Anfrage und versuchen sie so zu konzipieren, dass sie Teil von parsX wird und Ihnen somit nur ein Bruchteil der Entwicklungskosten entsteht.

## parsX – die Bausteine und Module

Um allen Erfordernissen der Verlagskunden hinsichtlich Verlagsprogramm, geplanten Prozessen und geplanten digitalen Produkten gerecht zu werden, haben wir parsX modular aufgebaut. Sie lizenzieren nur die parsX-Module, die Sie für Ihre Arbeit und Ihre Wunschprozesse benötigen. In jedem der Module finden Sie einen oder mehrere der unten beschriebenen parsX-Bausteine, die z.T. nicht einzeln lizenziert werden können.



### Die parsX-DTD

In jedem XML-Workflow ist die XML-Grammatik (die DTD) das zentrale Element. Hier wird festgelegt, in welchen Strukturen Ihre Daten künftig vorgehalten werden, was als korrekt und was als fehlerhaft angesehen werden soll. Die parsX-DTD ist der Branchenstandard für Publikumsverlage. Keine andere verfügbare DTD ist speziell auf die Bedürfnisse von Belletristik- und Sachbuchverlagen ausgelegt, die crossmedial publizieren wollen. In die überarbeitete Version 3.0 der DTD sind die Erfahrungen aus über 10.000 produzierten Titeln und viele Wünsche lizenznehmender Verlage eingeflossen. Auch künftige Weiterentwicklungen werden in enger Abstimmung mit der Anwenderschaft vorgenommen. Die parsX-DTD ist selbst auch wieder modular aufgebaut, so dass Sie immer nur die Elemente angeboten bekommen, die für Ihren Verlag relevant sind: So setzen Sie z.B. das Dramenmodul nur dann ein, wenn in Ihrem Verlag auch tatsächlich Bühnenstücke verlegt werden. Gleichzeitig ist mit dem Semantikmodul eine einfache, verlagsspezifische Erweiterung der DTD möglich, um Ihre ganz speziellen Strukturbedürfnisse abzubilden.

## Das parsX-Schematron

Valide XML-Daten gemäß DTD sind die wichtigste Voraussetzung für jede XML-Verarbeitung. Aber: Validität alleine reicht oft nicht aus. Denn jeder Verlag hat darüber hinaus eigene Anforderungen an seine Daten, die sich über ein DTD nicht abprüfen lassen, sogenannte *Konventionen*. Daher haben wir in parsX einen Schematron-Konventionsprüfer eingebaut, der bereits zahlreiche Plausibilitätsprüfungen standardmäßig durchführt, sich darüber hinaus aber auch leicht an Ihre spezifischen Bedürfnisse anpassen lässt.



### Der parsX-DokuChecker

Unseren DokuChecker werden Sie lieben. Er ist neben DTD und Schematron der dritte Qualitätssicherungsbaustein von parsX – und er verfolgt einen völlig anderen Ansatz. Der DokuChecker ist für XML das, was ein *preflight-check* für PDF-Daten ist: Eine Prüfung Ihrer Daten, ohne dass der Anwender Kenntnisse von XML haben muss. Abgeprüft und dokumentiert werden hier über einen Online-Dienst neben Plausibilitäten sämtliche strukturellen Besonderheiten der XML-Datei, jedes einzelne Zeichen, aber auch die von Ihnen gewünschte Mikrotypographie (Festabstände). Das Ergebnis ist eine PDF-Datei mit freisprachlichen Fehlermeldungen und Dokumentationen. Und da Ihr Dienstleister auf denselben Online-Dienst zugreifen kann, sieht er die Fehlermeldungen noch bevor Sie eine Eingangsprüfung der Daten machen und kann sie korrigieren.



### Das oXygen-Framework

Für die Bearbeitung Ihrer XML-Daten und als „Schaltzentrale“ für den Aufruf der Module setzen wir auf den bewährten XML-Editor *oXygen* unseres Technologiepartners SyncroSoft. Das parsX-Framework für *oXygen* macht aus dem unkonfigurierten Editor eine angepasste parsX-Oberfläche. Mit graphischer *wysiwym* („what you see is what you mean“)-Oberfläche und einfacher Schaltflächen-Steuerung für alle parsX-Module ist der *oXygen*-Editor Ihr Haupt-Arbeitswerkzeug. Auch Ihre E-Books und Leseproben werden künftig mit wenigen Mausklicks aus dem parsX-Framework im *oXygen* produziert.



### Das parsX-InDesign-PlugIn

Mit dem parsX-InDesign-PlugIn wird Ihr InDesign (und das Ihrer Dienstleister) zu einem Roundtripping-fähigen XML-Satzsystem für parsX-Daten. Eigene XML-Import- und -Export-Mechanismen und zahlreiche Erweiterungen der Funktionalität beheben die Schwächen, die InDesign standardmäßig im Umgang mit XML-Daten hat. Dank der mitgelieferten Muster-Templates, die den Leistungsumfang der parsX-DTD bereits aufnehmen, produzieren Sie im Satz so schnell wie nie zuvor – ohne dabei auf Teile des Leistungsumfangs von InDesign verzichten zu müssen.



### Der parsX-E-Book-Konverter

Neben dem InDesign-PlugIn ist der E-Book-Konverter das Herzstück von parsX. Mit knapp 200 Konfigurationsmöglichkeiten definieren Sie ein E-Book-Profil, das den Ansprüchen Ihres Verlages bis ins Detail entspricht: *Wo steht das Impressum? Wie lautet eine Titelei für eine Leseprobe? Wie viele Hierarchien soll das generierte Inhaltsverzeichnis enthalten? Sollen die Seitenwechsel aus dem gedruckten Buch in eine PageList überführt werden? Wie soll mit den Abbildungen und Tabellen umgegangen werden? Sollen Werbeseiten zugespielt werden?* Dieses Profil kombinieren Sie mit beliebigen, hinterlegbaren E-Book-Layouts im CSS-Format – und erstellen ab sofort auf Knopfdruck komplexeste E-Books (EPUB2, EPUB3, MobiPocket), die genau Ihren Vorstellungen entsprechen.

Eingebaute Qualitätssicherungstools wie der *pagina-GlyphenChecker* (prüft jedes Zeichen daraufhin ab, ob es in dem von Ihnen angegeben E-Book-Font vorhanden ist) und der *pagina-FontSubsetter* (erstellt zur Laufzeit der E-Book-Produktion einen eigenen Font, der nur die benötigten Zeichen enthält) stellen die maximale Qualität Ihrer E-Books sicher. Mit eigenem E-Book-Leseproben-Workflow.



#### Der parsX-PDF-Previewer und Leseproben-Konverter

Mit dem PDF-Previewer erzeugen Sie vollautomatisiert Voransichten Ihrer XML-Daten über den neuen "PrintCSS"-Standard oder XSL-FO. Genau wie beim E-Book-Konverter lassen sich beliebig viele Profile („Templates“) anlegen – nutzen Sie den PDF-Previewer zur Generierung spezifischer Ausgabeformen für die Bearbeitung im Lektorat, für eine rasche Umfangsschätzung oder um den Text in verschiedenen Reihenlayouts zu testen.

Ein kostenloser XSL-FO-Renderer ist bereits in der Basisversion enthalten. Die Steuerung des Satzbildes erfolgt dabei komfortabel über ein InDesign-Template – es war noch nie so einfach, XSL-FO zu konfigurieren!

Noch einfacher wird es mit PrintCSS. Mit unseren vorgefertigten PrintCSS-Stylesheet kann jeder, der die CSS-Syntax kennt, perfekte Buchlayouts erstellen – das daraus resultierende Satzbild hält (soweit es sich überhaupt automatisieren lässt) jedem Vergleich stand und wird von immer mehr Verlagen als günstige, vollautomatische Satzalternative eingesetzt – z.B., um Titel auch im Großdruck barrierefrei verfügbar zu haben oder um einen nur als E-Book verfügbaren Titel auch als *Print-On-Demand*-Titel anbieten zu können.

Mit eigenem Workflow für Print-Leseproben.

### Das Multimedia-Modul

Das Multimedia-Modul ist eine Erweiterung der DTD und des E-Book-Konverters, um multimediale Inhalte (Video, Audio, Bildergalerien) in EPUB-3-Dateien einbinden zu können. Erstellen Sie ein eigenes Profil für *enriched E-Books* und produzieren Sie diese genauso komfortabel und einfach wie Ihre normalen E-Books.

### Das Dramensatz-Modul

Mit dem Dramensatz-Modul können Belletristik-Verlage auch komplexe Bühnendialoge, Regieanweisungen etc. auszeichnen und über den E-Book-Konverter im E-Book ausgeben. Die Satzunterstützung ist in allen gängigen XML-fähigen Werksatzsystemen getestet; eine volle Integration in InDesign ist aufgrund technischer Limitierungen leider nicht möglich.

### Das Import-Modul

Mit dem Import-Modul, einer DTD- und EPUB-Konverter-Erweiterung, können andere Werke in parsX-XML-Struktur dateisystembasiert referenziert und im EPUB als Leseprobe angehängt werden. Vor jeder „Fremdleseprobe“ wird eine Leseproben-Titelei mit den wichtigsten Infos zur Leseprobe ergänzt. So können sehr einfach und komfortabel Leseproben-Empfehlungen an EPUB-Titel hinzugefügt werden.

## Die Datenbank-Schnittstelle / Kommandozeilen-Schnittstelle

Über die Datenbank-Schnittstelle lassen sich alle parsX-Module auch außerhalb des oXygen-Frameworks aufrufen und steuern. parsX lässt sich damit vollständig in Ihre bestehende oder künftige CMS/MAM-Umgebung integrieren.

Die Anwendung dieser Kommandozeilenbasierten Steuerung ist aber selbstverständlich nicht auf Datenbanken beschränkt, sondern erlaubt Ihnen zum Beispiel die EPUB- oder PDF-Erstellung auch im Backend Ihrer Website oder Ihres Webshops.

## Das Index-Modul

Funktionale, wissenschaftliche Register im E-Book waren bisher fast ein Ding der Unmöglichkeit. Das parsX-Index-Modul erlaubt Ihnen, komplexe (auch mehrstufige) Register in verschiedenen Varianten im E-Book abzubilden, z.B. auch in einer mediengerechten „*Keyword-in-Context*“-Darstellung mit einer Website-ähnlichen Index-Navigation im E-Book.

Aus den mit Registerbegriffen angereicherten parsX-XML-Daten lässt sich über das parsX InDesign-Plugin auch das native InDesign-Indexbedienfeld befüllen, so dass Sie auch Ihr Printwerk mit automatisch erstellten und verlinkten Registern versehen können. Die InDesign-Roundtripping-Funktionalität von parsX stellt zudem sicher, dass Änderungen am Index und den Indexbegriffen auch wieder ins E-Book zurückgespielt werden können.

## Der parsX-Serials-Generator

Der Serials-Generator enthält eine Erweiterung der DTD zur Auszeichnung von „Serials“ sowie ein oXygen-Plugin zum Splitting Ihrer Werke in einzelne Serials. Somit können Sie ein Werk automatisiert in verkaufbare Einzelkapitel oder andere, definierbare Gliederungseinheiten aufsplitten.



## Der pagina-Word-Konverter und das Word-Satzmodul

Der Word-Konverter ist kein integriertes Modul von parsX, sondern ein eigenständiges pagina-Produkt, das wir hier in einer auf die parsX-DTD vorkonfigurierten Fassung auch für parsX-Kunden anbieten. Es enthält die pagina-Wordtools, eine Sammlung von ca. 20 Word-Erweiterungen zum systematischen Copy-Editing, zum Bereinigen und Vorbereiten Ihrer Word-Daten als Grundlage für eine automatisierte Weiterverarbeitung. Der anschließende Konvertierungsprozess überführt Ihre Word-Daten nach parsX und auf Wunsch wieder zurück nach Word. Beim Rückspielen nach Word haben Sie zwei Optionen: „Enriched Word“ mit eingespielten Angaben zur Dokumentstruktur sowie das „Word-Satzmodul“, in dem Sie Autor und Lektorat in einer dem endgültigen Umbruch sehr angenäherten Anmutung arbeiten lassen können.

### Exkurs: XML-first oder XML-last?

*parsX schreibt Ihnen nicht vor, in welchem Prozess Sie Ihre Daten entstehen lassen. Es enthält alle erforderlichen Bausteine, um einen vollständigen XML-first-Workflow im Verlag komfortabel zu implementieren – aber das ist keine Bedingung. Je nach Verlagsbereich, vielleicht sogar je nach Mitarbeiter, Terminlage etc. entscheiden Sie, ob Sie ein Werk XML-first oder XML-last produzieren wollen. Ihre E-Book-Qualität wird immer dieselbe sein. Auch wenn Sie derzeit einen XML-last-Prozess mit anderen Partnern etablieren, ist parsX die ideale Ergänzung, um die E-Book- und Leseproben-Produktion zu standardisieren, die Qualitätssicherungsmechanismen zu nutzen und somit eine datentechnisch einheitliche Backlist aufzubauen. Die meisten parsX-Kunden entscheiden sich, mit einem XML-first-Workflow beim Dienstleister zu beginnen und die Softwaremodule im Verlag erst nach und nach (meist in der Herstellung, seltener im Lektorat) zu nutzen.*

*pagina empfiehlt einen XML-first-Workflow.*

## Kontakt

pagina GmbH Publikationstechnologien  
Herrenberger Straße 51  
72070 Tübingen

Tobias Fischer, Produktmanager parsX  
Tobias Ott, Geschäftsleitung

+49 (0) 7071 9876-44  
+49 (0) 7071 9876-54

tobias.fischer@pagina.gmbh  
tobias.ott@pagina.gmbh

[www.parsx.de](http://www.parsx.de)  
[www.pagina.gmbh](http://www.pagina.gmbh)