

parsX wird offener Herstellerstandard

Im Herbst 2019 veröffentlicht pagina die neue Version von parsX, dem leistungsfähigen und komfortablen Produktionssystem, mit dem sich Verlage konsequent auf die Anforderungen der digitalen Transformation ausrichten. Der Software zugrunde liegt ein XML-Datenstandard, der in enger Kooperation mit 19 Publikums- und Sachbuchverlagen entwickelt wurde und die spezifischen Bedarfe dieser Verlagssparte umfänglich abbildet. Es ist die einzige Wissensarchitektur ihrer Art für den deutschsprachigen Verlagsmarkt – und dieser Standard wird als freies XML-Dokumentenformat veröffentlicht, mit Hilfe eines Förderprogramms des Bundesministeriums für Wirtschaftliche Zusammenarbeit und Entwicklung.

Wieso ein offener XML-Herstellerstandard?

In Industrie und Wissenschaft hat sich längst die Erkenntnis durchgesetzt, dass der Einsatz branchenspezifischer Datenstandards der Schlüssel zum Erfolg in der digitalen Zeit ist. Die wichtigsten Standards – HTML, CSS, EPUB, BPMN etc. – sind allen geläufig, auf ihnen basiert der Erfolg eines „Publishing 4.0“. 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. Für die Anforderungen der Publikumsverlage gab es hingegen bislang keine erprobte und bewährte XML-Struktur, obwohl die Anforderungen täglich steigen. Solange aber jeder Verlag seine eigene Struktur entwickelt – oder sich eine solche Entwicklung gar nicht leisten kann –, haben die Publikumsverlage keine Lobby; es gibt keine schnell implementierbaren Softwarelösungen, keine kostengünstige Implementierung in CMS- oder MAM-Systeme, keine technischen Kooperationen.

parsX ist der einzige XML-Standard für Publikumsverlage im deutschsprachigen Raum. Er ist ausgereift und zehntausendfach erprobt. Eine immer größer werdende Verlagscommunity setzt parsX als zentrales Datenformat ein und stellt gemeinsam mit uns sicher, dass der Standard stets alle Anforderungen abdeckt, die ein moderner Verlag an seine Daten stellen kann.

Es war an der Zeit, aus diesen Erfahrungen ein offenes XML-Datenformat abzuleiten und der Verlagsbranche frei zugänglich zu machen. Wir wollen damit unseren Beitrag leisten, dass die digitale Transformation in der Verlagsbranche gelingen kann.

Der Aufwand hierfür ist erheblich, und so haben wir uns beim Bundesministerium für Wirtschaft um Förderung für dieses Projekt bemüht – und erhalten. Seither findet in enger Abstimmung mit den parsX-Bestandskunden ein komplettes Rebuild der XML-Grammatik statt sowie eine vollständige, umfangreiche Dokumentation.

Mit der neuen Version von parsX können wir somit der Verlagsbranche die modernste, umfangreichste und ausgereifteste XML-Grammatik anbieten, die je für Publikums- und Sachbuchverlage entwickelt wurde. Vollständig kostenlos!

Fragen und Antworten zur freien Veröffentlichung von parsX

1. Was genau muss man sich unter dem frei zugänglichen Schema vorstellen?

„Sobald man Daten austauschen – oder verarbeiten – möchte, stellt sich automatisch die Frage: In welchem Format sollen die Daten vorliegen? Man unterscheidet hierbei zwischen proprietären Datenformaten (Word, InDesign etc.) und offenen Datenformaten (HTML, XML etc.). Es herrscht Einigkeit darüber, dass nur ein offenes Datenformat wie XML eine langfristige Verfügbarkeit von (Verlags-)Inhalten sicherstellt. Doch damit ist noch nicht definiert, wie die Daten konkret strukturiert sein sollen – das übernimmt eine sogenannte XML-Grammatik, die meist in der Form einer DTD oder eines XML Schemas vorliegt. Hierin wird präzise geregelt, wie welche Dokumentstruktur ausgezeichnet sein muss, um dem Standard zu entsprechen. Auch hinter der „Internet-Sprache“ HTML steckt eine solche Grammatik. Nur weil alle Webseiten dieser Grammatik folgen, können Sie auf jedem Browser dargestellt werden.

Wir haben uns bei parsX für ein modulares XML-Schema entschieden. Es definiert alle Strukturen, die ein Publikums- und Sachbuchverlag benötigt, um seine Inhalte optimal aufzubereiten – für alle bekannten Medienformen, mit allen denkbaren Erweiterungen und Ausprägungen.

2. Gibt es nicht schon andere XML-Standards, die man verwenden kann?

„Es gibt zahlreiche Standards wie z.B. docbook oder DITA, oder eben auch HTML – aber alle diese Grammatiken bilden nicht die branchenspezifischen Anforderungen eines Verlages ab. Sie dienen z.B. der technischen Dokumentation. Das führt dazu, dass einerseits verlagsspezifische Elemente fehlen (beispielsweise für komplexe Registerstrukturen, für Multimedia-Elemente oder auch, um die Grenze einer Leseprobe zu definieren), andererseits beinhalten sie zahlreiche Elemente, die für Verlage überhaupt keine Rolle spielen.

3. Was macht die parsX-Struktur so besonders für unsere Branche?

„parsX wird mit dem Anspruch entwickelt, alle verlagsspezifischen Strukturen von Texten abzubilden – von eingekürzten Kolumntiteln bis zu Schmuckinitialen, von Mehr-autorenwerken bis zu Lyrik und Bühnenstücken. Auch sind sämtliche XML-Strukturen in deutscher Sprache verfasst, um eine klare und eindeutige Branchensprache zwischen Verlagen und Dienstleistern sicherzustellen und eine intuitive Bedienung zu gewährleisten. Die Branche hat sich viel zu lange damit beholfen, andere Standards für die eigenen Anforderungen umzubauen und zweckzuentfremden. Das wollen wir ändern.

4. Umfasst die kostenlose Version den gesamten Leistungsumfang der XML-Grammatik? Oder ist das ein „parsX light“?

„Die gesamte XML-Grammatik wird kostenfrei veröffentlicht, ebenso die gesamte Dokumentation, zusammen mit Übungswerken und Tutorials. Es gibt keine Unterscheidung in eine kostenlose und eine kostenpflichtige Version. Auch sämtliche Möglichkeiten, die Grammatik an die eigenen Verlagsbedürfnisse anzupassen („Semantik-Modul“) werden kostenlos veröffentlicht.

5. Wird es weiterhin eine kostenpflichtige Version von parsX geben?

„ Ja, die einzelnen Softwaremodule, die auf dem XML-Standard aufsetzen, werden von pagina weiterentwickelt und sind weiterhin als lizenzpflichtige Software erhältlich. Das sind insbesondere die Module für eine automatisierte EPUB-Produktion, die vollautomatische PDF-Produktion über PrintCSS sowie das InDesign-PlugIn für eine individualisierte Satzproduktion. Aber auch jeder andere Softwareentwickler kann künftig auf der Basis des Standards entsprechende Publikationssoftware anbieten. Schon heute gibt es zahlreiche Kooperationen in dieser Richtung.

Stand: September 2019

pagina GmbH Publikationstechnologien
Herrenberger Straße 51
72070 Tübingen
www.pagina.gmbh www.parsX.de

Tobias Ott (Geschäftsleitung)
tobias.ott@pagina.gmbh
07071-987654

- gegründet 1973
- 30 MitarbeiterInnen
- Geschäftsbereiche: XML-Medienproduktion, Softwareentwicklung,
Digitale Bildung, digital humanities