

Exakter Bauplan

XML Die medienneutrale Datenhaltung gilt als Wunderwaffe bei der rationellen Verarbeitung von Inhalten. Funktionieren kann sie aber nur, wenn die richtigen Strukturen geschaffen werden. Ein Werkstattbericht aus dem Verlag Walter de Gruyter.

Text: Ursula Welsch
Mitarbeit: Martina Bach
(de Gruyter),
Ingrid Schmidt
(Fraunhofer IPSI)

Wer den neuen »Psychyrembel. Klinisches Wörterbuch« (260. Auflage) zur Hand nimmt, stellt rein äußerlich keinen Unterschied zur Vorgängerauflage fest. Dabei hat sich der Produktionsprozess fundamental gewandelt. Vor zwei Jahren hat der Berliner Wissenschaftsverlag Walter de Gruyter seinen Kerntitel im medizinischen Programm – und die darum herum gewachsene Produktfamilie – auf die medienneutrale Datenhaltung umgestellt. Zugrunde lag der Wunsch, die Produktion auf der Basis standardisierter Produktionsprozesse zu optimieren – und zwar ohne inhaltlichen Qua-

litätsverlust. Weiteres Ziel: Inhalte sollten technisch so aufbereitet werden, dass sie – ohne zusätzlichen Aufwand – in verschiedenen elektronischen Produktformen verwertet werden können. Eine Aufbereitung in XML bot die Möglichkeit, beide Ziele zu erreichen. Dass dabei auch eine durchgängige formale Konsistenz für alle Wörterbücher der »Psychyrembel«-Familie erreicht werden konnte, war ein gern akzeptierter Zusatzeffekt.

Projektorganisation

Wie geht man ein so wichtiges und umfangreiches Projekt an?

Zunächst wurde unter der Leitung der Chefredaktion ein Projektteam zusammengestellt – mit der Aufgabe, eine Dokumenttyp-Definition (DTD), das heißt eine maschinenlesbare Strukturbeschreibung der Inhalte, in XML zu entwickeln. Die DTD sollte sowohl die Vorstellungen des Verlags über die künftige Nutzung der Substanzen als auch die redaktionellen Abläufe und die Gegebenheiten des Content-Managementsystems (CMS) berücksichtigen. Daher setzte sich das Team aus Mitgliedern der Redaktion, den DTD-Entwicklerinnen und einer Vertreterin des Dienstleisters für das CMS zusammen.

In der Anfangsphase traf sich das Team ungefähr einmal im Monat, später reichten Treffen im Turnus von zwei Monaten aus.

Begonnen wurde die DTD-Entwicklung mit dem Hauptwerk, »Psyrembel Klinisches Wörterbuch«, da es inhaltlich das Zentrum der Wörterbuchfamilie bildet. Als das Grundgerüst der DTD dafür stand, wurde als zweites Produkt der »Hunnius«, Pharmazeutisches Wörterbuch, in die Analyse einbezogen.

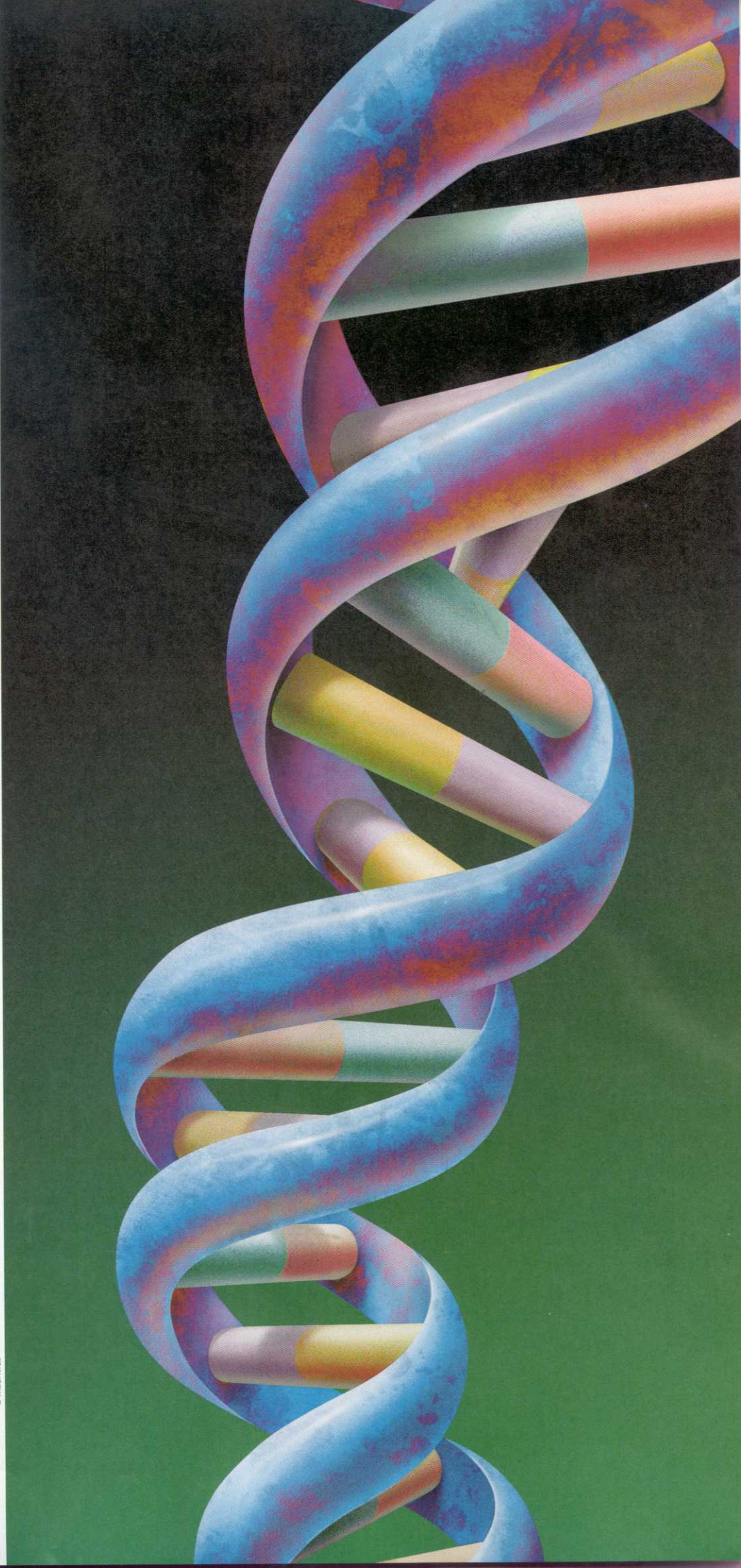
DTD-Erstellung

Basisanforderung an die DTD-Erstellung war, dass aus den strukturierten medienneutralen Daten sowohl die Printfassung als auch elektronische Ausgaben optimal hergestellt werden können. Zielvorgabe war außerdem, dass eine gemeinsame DTD für die gesamte Wörterbuchfamilie entstehen sollte. Im Lauf der Treffen haben sich weitere Ziele herauskristallisiert, die in die DTD-Entwicklung eingingen:

■ Die DTD des »Klinischen Wörterbuchs« dient als Leitstruktur für die übrigen Wörterbücher. Die Inhalte aller Wörterbücher der »Psyrembel«-Familie sollen im Rahmen der Online-Präsentation gemeinsam recherchierbar sein.

■ Die XML-Struktur orientiert sich an der Bedeutung der medizinischen und pharmazeutischen Inhalte. Damit war eine klare Ausrichtung auf eine werkneutrale, semantische Strukturierung gegeben.

Mit jedem weiteren Wörterbuch, das einbezogen wurde, zeigte sich, dass die einzelnen Titel aufgrund ihrer unterschiedlichen thematischen Schwerpunkte sehr individuell aufgebaut sind, so dass eine vorab entwickelte identische Struktur für alle Wörterbücher der »Psyrembel«-Familie unweigerlich mit einem Qualitätsverlust bei den einzel- ➤



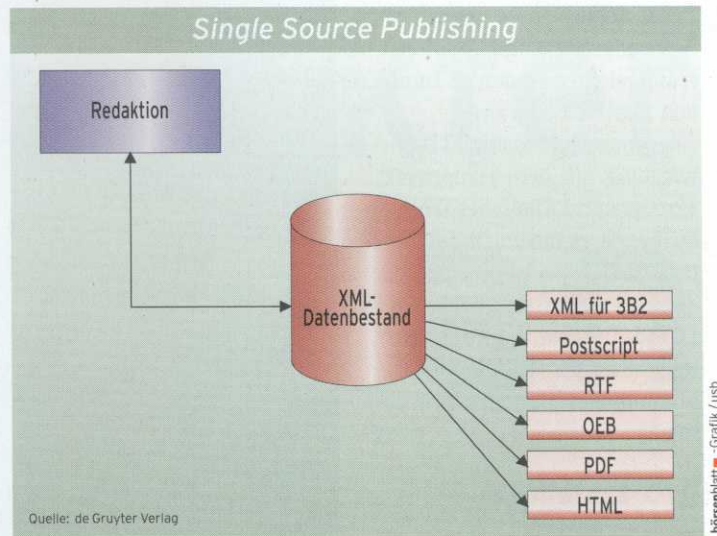
➤ nen Substanzen einhergehen würde. Vor diesem Hintergrund fiel die Entscheidung, semantische DTD-Module zu entwickeln, aus denen sich die Einzel-DTDs flexibel zusammensetzen lassen.

Bei der Konvertierung der medizinischen Wörterbücher konnte die sehr komplexe, aber in hohem Maß formal konsistente lexikografische Verdichtung aller in der internen Datenbank gespeicherten Daten genutzt werden. Der »Hunnius« hingegen lag in Form von extern erarbeiteten Word-Daten vor. Aufgrund der semantischen DTD-Konzeption wurden bei der Datenübernahme des »Hunnius« differierende lexikografische Festlegungen gezielt durch diejenigen des »Klinischen Wörterbuchs« ersetzt. So konnte der »Hunnius« im Zuge der Neustrukturierung gleichzeitig formal auf das Niveau der anderen »Psyhyrembel«-Wörterbücher gehoben werden.

Die Datenmigration in das reichhaltige XML-Format war dennoch nicht immer einfach. Eine sorgfältige redaktionelle Nachbearbeitung war notwendig. Sie wurde parallel zu inhaltlichen Überarbeitungen durchgeführt. Während des Arbeitsprozesses ergaben sich immer wieder Situationen, die eine Fortschreibung der DTDs notwendig machten, etwa weil weitere Titel in den Umstellungsprozess einbezogen wurden. Das zugrunde gelegte Modulsystem der DTDs sorgt dafür, dass sich solche »spontanen« Anpassungen in den systematischen Charakter der DTDs einfügen.

Heutiger Stand

Parallel zur redaktionellen Aktualisierung der ersten beiden Werke wurden vier weitere Wörterbücher der »Psyhyrembel«-Familie in das XML-Konzept integriert und in das CMS eingespielt. Die Arbeitstreffen haben sich unterdessen auf einen Tag je neu ein-



Quelle für viele Anwendungen:
XML-Daten können in verschiedene Formate umgewandelt werden

zufügendem Wörterbuch reduziert. Durch die gezielten Maßnahmen in der DTD konnten die Durchlaufzeiten bei der Bearbeitung verkürzt werden, wodurch ein früherer Return-on-Investment möglich wird, als es bei einem konventionellen Vorgehen der Fall wäre.

Konkrete Kostensenkungen für den Satz werden im Lauf der folgenden Produktionen erzielt – unabhängig davon, ob es sich dabei um die Herstellung eines weiteren Wörterbuchs der »Psyhyrembel«-Familie handelt oder um eine Neuauflage des »Klinischen Wörterbuchs« –, denn die vorliegende Satzprogrammierung muss nur geringfügig angepasst werden.

Die Umstellung der Produktion auf die medienneutrale Datenhaltung mit XML hat den ersten Praxistest durchlaufen und kann – bei allen Investitionen an Zeit und Geld, die ein neues Redaktions- und Produktionssystem zu Beginn erfordert – aus Sicht der Beteiligten als Erfolg gewertet werden. Die Grundkonzeption der Wörterbuch-beziehungsweise Artikelstruktur ist auch auf Wörterbü-

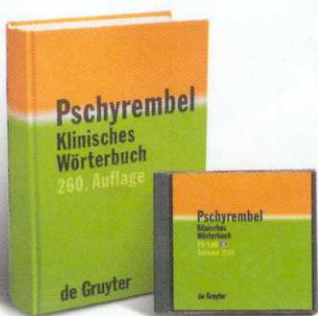
cher anderer Fachbereiche übertragbar. Darüber hinaus kann über zusätzliche Nutzungen aus der Substanz nachgedacht werden, zum Beispiel über ein deutsch-englisches Glossar, das aufgrund der semantischen Strukturierung automatisiert aus dem Datenbestand aller Wörterbücher der »Psyhyrembel«-Familie erzeugt werden kann.

Die nächsten Ziele stehen bereits fest – die wichtigsten Aufgaben sind:

- E-Book-Publikation des »Klinischen Wörterbuchs«
- Integration der bereits migrierten Datenbestände in das Online-Portal www.psyhyrembel.de
- Strukturierung weiterer Wörterbücher der »Psyhyrembel«-Familie.

Auch die Optimierung des Workflows – vom Datenaustausch mit den Autoren über die Textbearbeitung bis hin zur Generierung der unterschiedlichen, in Zukunft möglicherweise stark individualisierten Verkaufsformen (E-Book, Intranet etc.) – wird weiterhin im Fokus bleiben. Das Potenzial, diese Ziele zu erreichen, ist vorhanden. XML dient als Schlüssel dazu. **b**

Die Autorin ist freie Beraterin, Gründerin der XML-Schule und Dozentin für elektronisches Publizieren



Zwei Medien, ein Inhalt:
Beim »Pschyrembel« wurde die medienneutrale Datenhaltung umgesetzt