JATSを含むXML運用のコツ

- XMLは出版業界における最大のDX -

2022年10月12日アリエッタ株式会社山岡裕幸・佐藤秀幸

XMLでできること・メリット

自動組版 (InDesignやTeX等の組版ソフトを使用) および情報処理によって下記が実現可能。

- ▶ 大量の情報かつ定形フォーマットの書籍を効率的に制作
- ► XMLはHTMLとの相性が良いためWEBコンテンツを効率的に制作

XMLを使用した製作例

> 学術論文

規格: JATS XML

生成物:印刷用PDF, WEB用PDF(リンク付), J-Stage公開用全文XML

▶ 企業の製品マニュアル

規 格:**独自構造XML**

生成物:印刷用PDF, WEB用PDF (リンク付), オンラインマニュアル用HTML, CHM (ヘルプファイル)

▶ 塾・予備校などの模擬試験問題冊子+解説書

規 格:**独自構造**XML

生成物:印刷用PDF, WEB用PDF (リンク付)

XMLで作る際のデメリット

▶ 技術面でもハンドリング(ディレクション)面でも難易度が高い。

優秀なオペレーターはもちろん,技術をある程度理解しつつ,さらに顧客業務や案件の特徴を理解しながらお客様との折衝 (仕様検討・料金設定など) ができる人材も必要。

ト 赤字修正が大変

大量のタグが付与されたソース (テキストデータ) を修正するので, 修正量が多いものは 不向き。→**運用に工夫が必要**

理想的なXMLの運用方法

XMLを使用して自動組版し、「Single Source Publishing」「ワンソースマルチユース」, つまり1つのデータから色々なデータを自動生成する。

- ▶ 印刷用PDF
- ▶ WEB用PDF(各種リンク付)
- HTML
- その他形式のデータ

※XMLはあくまで言語であって組版ソフトではないので、InDesign等の組版ソフトが別途必要。

※XMLの前にExcel/FileMaker/Access等のソフトを使用して,これらのデータを"ワンソース"とする(した方が良い)こともある。

理想通りに行かない時

- ▶ 赤字修正が多い
- ▶ 既存の印刷・制作会社では対応できないが変えられない(変えたくない)
- その他もろもろの事情

そのような場合は,

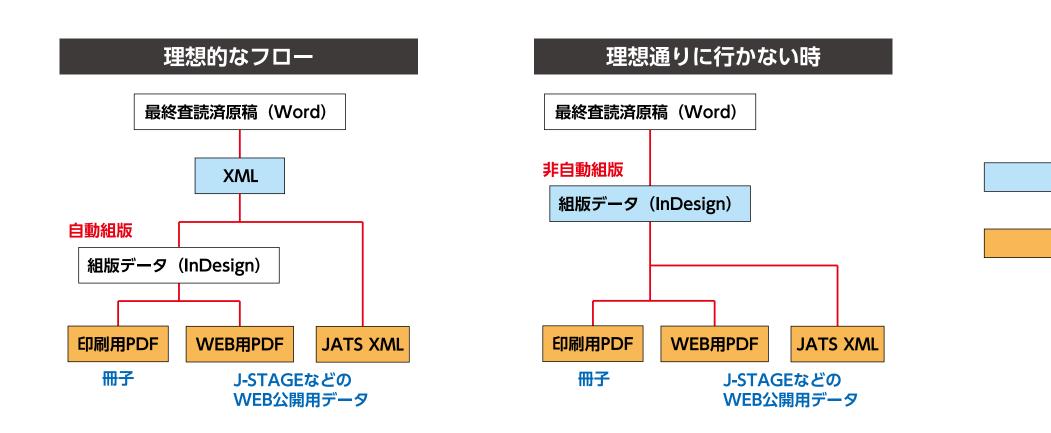
組版データ (InDesignを想定) からのXMLデータを作成 (JATS XML含む)

※理想通りの運用方法でないと実現できない場合もあります。

マスターとなるデータ

生成データ

運用フロー例(学術論文の場合)



組版データからXMLを作成する際のコツ・注意点①

「大は小を兼ねる」

細分化された情報をまとめることは簡単だが、その逆は困難(非常に手間がかかる)。

例) 姓名のデータ

姓	名	─	姓名
蛇津	太郎	X	蛇津太郎

つまり最初から**細分化されたデータで管理**すれば、InDesignデータからも比較的容易にXMLに書き出すことが可能。

InDesignでは「**段落スタイル」「文字スタイル」**というレイアウト設定があり、その設定情報が含まれた「タグ付きテキスト」を書き出すことが可能。その「**タグ付きテキスト」をXMLに変換**すれば良い。

組版データからXMLを作成する際のコツ・注意点②

「組版データに意味を持たせる」

印刷・制作会社の組版(DTP)部門ではInDesignの場合,一般的に先述の「段落スタイル」「文字スタイル」の設定をするが,その際に見た目ではなく意味を持たせることがXML書き出しのコツとなる。例えば学術論文の場合,

例1) 著者名と所属のレイアウトが同じだとしても,「段落スタイル」は別々にする。

 段落スタイル名:
 NG
 13級20歯
 —

 OK
 著者名
 所属

▶ 例2) 所属番号の上付と本文中文献番号のレイアウトが同じだとしても,「文字スタイル」は別々にする。

 文字スタイル名:
 NG
 13級上付
 –

 OK
 所属番号
 本文中文献番号

JATS XMLの作成方法

【Wordから生成】

- ▶ eXtyles® (Inera社; Wordのアドイン) を使用※
- ▶ J-Stageが提供する全文XML作成ツールを使用
- ▶ Wordから生成する処理を自作
- ※一般的にはカスタマイズが必要。eXtyles®には文献チェック機能があるため便利。

【組版データから生成】

▶ テキストデータ (先述のInDesignの「タグ付きテキストデータ」がベター)を書き出して、 それに処理を加えてから作成 (一発で作成は極めて困難)

さいごに

- ▶ 効率的かつ効果的に出版・情報を伝達することが目的であり、その手段として (JATS XMLを含めて) XMLが存在する。
- ▶ 全てを自動組版・自動処理しようとすると逆に非効率(非現実的)なことがしばしばある。100 or ゼロで考えるのではなく,何かを妥協しつつ調整しながら仕様を決めて進めていくことが重要。

でも、確実に言えることは

「XMLは出版業界における

最大のDX (デジタルトランスフォーメーション) になりうる」

お問い合わせ

ご清聴いただきありがとうございました。



アリエッタ株式会社

代表取締役 山岡 裕幸

〒116-0013 東京都荒川区西日暮里6-41-5 酒正ビル203号

TEL: 03-6807-7693

Mail: hyamaoka@arietta.co.jp

URL: https://www.arietta.co.jp/

