
高品質な書籍を簡単に制作するための出版支援ソフトウェアの実用化

IdeoType: A Book Compiler

1. 背景

本プロジェクトでは、高品質な書籍を簡単に制作するための出版支援ソフトウェア”IdeoType”を開発する。IdeoTypeは、原稿からその場で自動的に本を作り上げることにより、書き手が試行錯誤により内容を改善することを支援し、高品質な書籍を短期間に低コストで制作することを可能にする。

現在日本語の商業出版で組版に使われているMac DTPソフトウェアは、非対話的に実行(バッチ処理)できない。一方バッチ処理組版ソフトウェアには、使いやすさが不足している。非対話的に実行できてかつ使いやすいソフトウェアがないため、著者が書き直しを繰り返して品質を高めることが困難になっている。

この状況を解決するには、それぞれの長所を組み合わせ、つまり「非対話的で」かつ「使いやすい」ソフトウェアが必要となる。IdeoTypeはそれを実現する。

2. 目的

本プロジェクトでは、IPA 2006年度上期未踏ソフトウェア創造事業「高品質な書籍を簡単に制作するための出版支援ソフトウェアの開発」において開発されたプロトタイプを発展させ、実用に耐えるソフトウェアにすることを目指す。2006年のプロトタイプでは、最終的に良い感触は得たものの、とにかく動作するものを作り上げることを優先し、拡張性などは後回しにしたため、成果物の実用性には改善の余地があった。本プロジェクトでは、プロトタイプで実現した「非対話的」「使いやすい」という特長を維持しつつ実用に耐えるソフトウェアに仕上げ、最終的には現場に投入することを目標とする。

3. 開発の内容

IdeoTypeは大きく分けて次の部分からなる。

- 変換ライブラリ

テキスト原稿をバックエンド用の形式に変換する機能。原稿形式にXHTML、PDF出力用のバックエンドにLaTeXを使っているため、html2latex.xslというXSLTスタイルシートを中心にした形での実装となっている。プロジェクトごとの設定で上書きすることが可能で、それによりカスタマイズが可能になっている。

- ユーティリティライブラリ

変換ライブラリを補助する機能群(ファイルの結合・正規化・数式記法対応など)。変換ライブラリから補助的な機能をユーティリティライブラリに積極的に分離することにより、主たる部分を単純に保ち、開発効率を維持する。RubyやXSLTで実装される。

- ビルドライブラリ

ライブラリ全体を統括するビルドスクリプト。分散していたライブラリ呼び出しをこのビルドライブラリに集約し、アプリケーション全体の開発を効率化する

- ユーザインタフェース
ライブラリを利用するためのユーザインタフェース。Rubyで記述。

上記以外の主要な構成要素としては、次のものがある。

- Tests
IdeoTypeはほぼすべてのプログラムにユニットテストが付属する(XSLTはxsltunit、RubyはTest::Unitを用いて記述)。これによりデバッグに割く時間を減らし、ソフトウェアを常に動作する状態に維持しながら開発することが可能となった。また、ユーザはテストを読めば、各機能がどのような動作を想定して実装されたのかを知ることができる。テストコードおよびテストデータはすべてtest/ディレクトリに収められ、test/Makefileのtestターゲットにより実行される。
- Documents
ドキュメントはdoc/ディレクトリに収められている。ソースはIdeoType自身で処理可能なXHTML+MathML形式で記述され、Makefileのdocターゲットにより閲覧に適した言語ごとのXHTMLファイルが生成される。WWW向けのドキュメントはwwwターゲットにより生成される。
- Makefile
テストを実行したり、ドキュメントを生成したり、IdeoType自身をパッケージングしたりといった開発者向けのタスクはMakefileに収められている。

次の点が主な特徴といえる。

- 商業出版で利用可能な実用性
商業出版の現場で利用可能なソフトウェアを開発した。
- 特殊な機材が必要ない
家庭用のPCで実行可能。Windowsではまだ動作しないが、VMware等でUnix環境を用意してそこにセットアップするのは、慣れれば1時間もかからない。
- オープンソースソフトウェアである
誰でも自由に利用可能。今後も継続してオープンな形で開発を続ける。

4. 従来の技術(または機能)との相違

既存の類似ソフトウェアと比べると、次のような特徴がある。いずれもXHTMLを原稿形式に採用したことに大きく依っている。

- 原稿データの可搬性・相互運用性・拡張性が高い
HTMLを媒介にして、ユーザは任意のソフトウェアを組み合わせることができる。BlogやWikiのバックエンドに使うことも可能。現時点で、ワープロ・Officeスイート・Blog・Wikiなど文書を扱うソフトウェアのほとんどがHTML形式で出力できることを考えると、入力に関しては最も広い受け口を持っているといえる。また標準に準拠しているため相互運用性は高く、XHTML+MathMLという形でXML由来の拡張性の高さも享受している。
- ユーザビリティが比較的高い
LaTeXマクロやXMLに比べてHTMLはルールが単純で簡単に記述できる。また教材やWebブラウザなどの関連ツールが豊富にあるため、知識や道具も流用しやすい。その結果、ユーザに

とって使いやすいものとなっている。

5. 期待される効果

本ソフトウェアにより、従来の出版業よりも高品質な出版物を効率良く作り出せる。制作作業を自動化することで、人的資源をレビューや推敲に集中できる。また、制作期間を短縮し、人件費を削減できる。品質を定量的に評価することは難しいので措くとしても、制作期間に関して言えば、本ソフトウェアを利用することで書籍企画1件につき数週間から数ヶ月短縮し、数十人時間分のコストを削減することが可能になる。

また、同一の内容を紙(PDF)とWebやCD-ROM(HTML, PDF)で出版するといった、Webと従来の媒体が融合した、媒体を問わない出版事業がローコストで可能になる。

本ソフトウェアはオープンソースソフトウェアであり自由に利用できるため、独占的なライセンスで提供した場合と比較してより広く使われることが期待される。

6. 普及(または活用)の見通し

本ソフトウェアを直接利用するユーザとしては、初期的に、出版を生業にしている職業編集者と、専門書の著者(技術者や研究者)を想定している。出版社の数や出版点数などから推測して、数年かけて数百人程度のユーザを獲得できれば、ある程度成功したといえるだろう。それだけのユーザがいれば間接的に影響する読者の数は年間で数十万人を超える。

なお、開発者が知る範囲では、商業出版を目的として試用中のユーザを2グループ、それ以外の目的で試用中のユーザを2グループ確認している(2008年3月現在。オープンソースソフトウェアとして配布しているため実際のユーザ数はそれ以上)。

最終的にはWebベースのUIを実装し、Blogツールの出力バックエンドなどとして広く一般に使ってもらうことを目標としている。ユーザ数は予想できないが、全世界で数万人以上に使ってもらえることを目指したい。

7. 開発者名

森田 尚 (hisashim@users.sourceforge.net, <http://sourceforge.net/projects/ideotype/>)