

オランダ「CB 社」のプレゼン資料の補足説明

CB 社は、元々出版社と実書店とで作られたジョイント会社なので、書籍サプライチェーンという両社の間は互角の関係にあり、オランダ市場の書籍流通（物流、商流）のセンターです。ジョイント会社として、CB 社はどちらにも有利に働くことはなく、サプライチェーンのすべてのために活動します。つまり、紙版の本、電子書籍・オーディオブック、両方の配送のための役目を持っているのでオンライン書店だけでなく実書店とともに活動しています。CB 社は顧客への直接販売は行なっていませんが、書店に一冊注文サービスという顧客への直接配送を提供しています。CB 社は紙版と電子書籍と関連商品のディストリビューター（販売流通代理店、代行業務に対して手数料を取る）ですが、(EDiTEUR の知る限り)取次業者としては活動していません。(CB 社の自社在庫ではないし、書店との間の売買には関与しない)

CB 社では、出版社向けに書籍の分類情報を更新するための Web ページ（スライド 16）のようなインターフェースを持っているようです。それは、最初のアルゴリズムによって基本的な推奨 Thema コードを見せて、出版社がそれを変更できるようになっているようです。既存の基本的な書籍メタデータを基にして、可能性のある Thema コードを推奨していますが、EDiTEUR の検索ブラウザからも選択できるようになっています。その推奨を行なっているアルゴリズムは、タイトル、既存の分類情報、シリーズ名、関連キーワードなど既存のメタデータフィールドのデータを見ています。それらは、フルテキストにアクセスして共通の単語や話題を探すなどの、いくつかの自然言語処理のようです。しかし、スライド 23 を見ると、そのアルゴリズムはまだ欠点があるようです。スライド 23 は、実際に自動で Thema コードを 1 つ割り当てる実例ですが、既存のメタデータの質によって同じ ISBN の本に色々な結果が現れています。例えば、書籍「DE MEESTE MENSEN DEUGEN（ほとんどの人は良い人）」は、英語翻訳版は、「Humankind:A Hopeful History」ですが、その推奨コードを見ると、ほとんどは、この本を分類する正しいコードのようには見えません。おそらく、本の目次、既存コードもしくは本の概要文を基本にしたアルゴリズムがコードを推奨しているのかもしれませんが。

CB 社は外部のデータ専門家を雇って、初期のアルゴリズムを作るサポートをしているようですが、企業の機密情報なので、基本的な概要のみ公開です。スライド 17 では、BOL.com 以外の 2 つの書店サイトでの Thema コードの試作運用を行っていたようですが、その結果については共有されていません。

<https://www.scheltema.nl/piedaterre/>

<https://www.donner.nl/>

これらの書店サイトは、CB 社のものではなく、CB 社が配送している書店です。この最初の実験では、タイトルに 1 つの Thema コードしか追加されていませんでしたので、複数の Thema サブジェクトコードとクオリファイアコードが提供される可能性については、このプレゼンでは検討されていませんでした。また、クオリファイアコードの追加、クロス検索、フィルタリング検索、ファセット検索、複数の適切なセクションにタイトルを置くなどの可能性についても検討していません。しかし、スライド 25 は、オランダ市場での Thema の将来的な利用の可能性を見ています。

(注 1) この補足説明は、EDiTEUR の担当者 Chris Saynor 氏からのメールを元に解説している。

(注 2) このプレゼンテーション資料は、2019 年 10 月に行われたもので、Bol.com はその後もプログラムの改善を重ねているが、数量など具体的な成果については残念ながら公表されていない。