

日本における電子書籍化の現状（2020年版）

—— 国立国会図書館所蔵資料の電子化率調査 ——

鷹野凌（NP0法人HON.jp）／共同研究者：堀正岳

概要

国立国会図書館所蔵資料の全件データと、電子書店「BOOK☆WALKER」の全商品データを、ISBNをキーにしてマッチングし、出版年別電子書籍化率などを調査した。ISBNのある国内出版物の電子書籍化率は、11.9%という結果が得られた。出版年別では、2017年29.6%、2018年31.2%、2019年33.2%と、徐々に電子書籍化率は高まっている。

キーワード：電子出版、電子書籍、電子書籍化、出版点数、国立国会図書館

1. 研究目的

出版科学研究所によると2019年の電子出版市場は、電子コミックが2593億円（同29.5%増）と急増しているⁱ。しかし、電子書籍（文字もの）は349億円（同8.7%増）とまだ市場が小さく、伸びも緩やかだ。この違いは、電子化点数や電子化率の差にあると考える。そこで、安形輝と上田修一（日本図書館情報学会）の先行研究ⁱⁱ（以下、安形／上田論文）を踏まえ、印刷本のうち電子書籍版が提供される割合の最新状況把握と、ランダムサンプリングではなく全件を対象とした、より詳細な分析を行う。

2. 研究手法

印刷本の書誌は、安形／上田論文と同様、国立国会図書館が収集している国内出版物のものをを用いた。国立国会図書館へ依頼したところ、書誌全件ファイル DC-NDL（RDF）形式（2020年1月29日時点、zip圧縮状態で3.87GB：以下、NDLデータ）が入手できたⁱⁱⁱ。NDLデータのうち、ISBNが存在するのは3,284,996件。そのうち、language値がjpnで年号が存在する2,623,535件

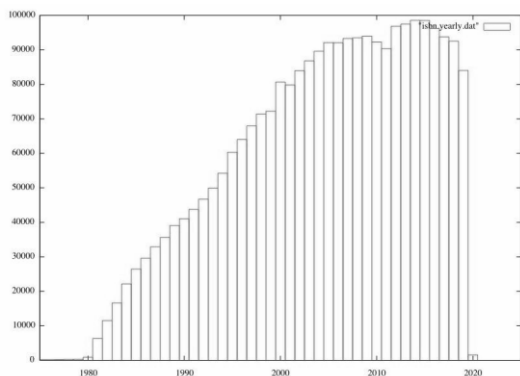


図 1. NDLデータ 出版年毎点数推移

を、マッチング対象とした。図1は、出版年別点数の推移グラフ。

このNDLデータのISBNをキーとして、電子書籍の書誌データをマッチングする。ただ、NDLデータはISBNが10桁時代（2006年以前）のものは10桁のまま収録されているため、先頭に978を付けてチェックデジットを再計算、13桁へ変換した。データ変換やマッチング処理にはPythonを用いた^{iv}。

電子書籍の書誌は、一般ユーザーが誰でもアクセス可能な電子書店^v「BOOK☆WALKER」が提供している、全商品RSS^{vi}（2020年2

月15日時点、1.54GB：以下、BWデータ）を用いた。「BOOK☆WALKER」は2010年12月にオープンした電子書店で、運営はKADOKAWAグループの株式会社ブックウォーカー。オープン当初はKADOKAWAグループの電子書籍だけを配信していたが、2012年から方針を転換、2013年から本格的に他社の電子書籍も配信するようになった^{vii}。現在は出版社横断型の、書籍もマンガも扱う一般的な総合型電子書店である。BWデータのうち、ISBNが存在するのは357,484件。図2はISBN有無別／配信年別の点数を推移グラフにしたもの。

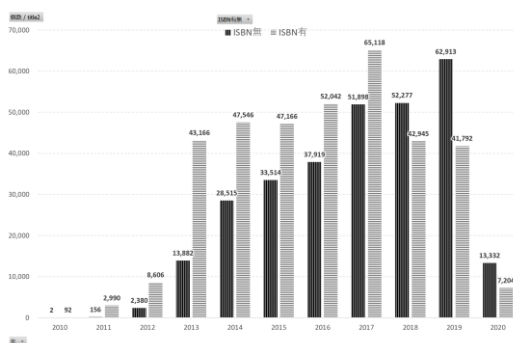


図 2 BWデータ ISBN有無別／配信年毎点数推移

BWデータのうち、NDLデータにISBNでマッチしたのは313,120件（87.59%）だった。NDLデータは印刷本の出版年だが、BWデータにあるのは電子書籍版の配信年のみ。そこで、電子書籍の底本がいつ出版されたかについては、ISBNでマッチングした上で、NDLデータを参照した。これにより、出版年毎の電子書籍化率が算出できた。なお、安形／上田論文は、Amazon Kindle、紀伊國屋書店、Google Books、Maruzen eBook Libraryを対象とした、各年5000件のランダムサンプリング。当調査の対象はBOOK☆WALKERのみだが、全件が対象。

3. 結果の知見

NDLデータのISBN有・日本語・年号有2,623,535件に対し、マッチしたBWデータは313,120件、電子書籍化率は11.9%であった（安形／上田論文では、全体の電子書籍化率が明らかにされていない）。安形／上田論文では出版年が2017年のタイトルは電子書籍化率が36.8%だが、当調査では29.6%だった。この差は、Amazon Kindle独占で配信しているケースや、学術系でMaruzen eBook Libraryのみで配信しているケースが考えられる。また、安形／上田論文の時点では未調査だった2018年の電子書籍化率は31.2%、2019年の電子書籍化率は33.2%と、出版年を追うごとにさらに電子書籍化率は高まっている傾向が見られた。図3は、NDLデータとBWデータをマッチングして得られた、出版年毎の電子書籍化率。



図 3 出版年毎の電子書籍化率

BWデータには「ジャンル」が存在するため、NDLデータとISBNでマッチングすることによって得られた底本出版年を用い、ジャンル別&底本出版年別の電子書籍点数をグラフ化した(図4)。マンガは、印刷本の出版年が古いものも電子書籍化された点数が比較的多いが、他ジャンルは少ないことがわかる。また、2010年以降については、文芸・小説・ライトノベルより、実用ジャンルのほうが点数的には多くなっていることがわかった。

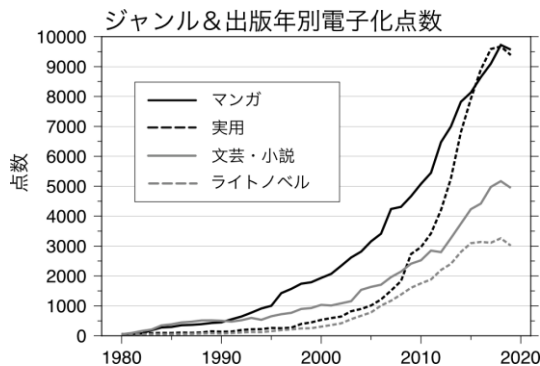


図 4 BWデータのジャンル&出版年別点数

本手法により、電子書籍化の実態を従来より精緻に把握できることがわかった。ただ、ジャンル別での集計や、出版社別での集計は、課題が残った。NDLデータのNDC分類やNDLC分類は、存在/不存在がバラバラなため集計に使えない。Cコードが存在しないため、安形/上田論文の手法と同様、他のデータからマッチングする必要がある。出版社名はリテラルで収録されており、表記揺れもあるため集計に使えない。ISBNに含まれる出版社記号(桁数が異なる)での集計手法を確立する必要がある。

-
- ⁱ 出版科学研究所「2019年出版市場（紙＋電子）を発表しました」
<https://www.ajpea.or.jp/information/20200124/index.html>
- ⁱⁱ 日本図書館情報学会誌 65巻（2019年）2号掲載「日本における電子書籍化の現状」（安形輝、上田修一）
https://doi.org/10.20651/jslis.65.2_84
- ⁱⁱⁱ 国立国会図書館「書誌情報提供サービス」
https://www.ndl.go.jp/jp/data/data_service/index.html
- ^{iv} ソースコードは Github `ndl-read/01parse_ndl.py` (mehori)
https://github.com/mehori/ndl-read/blob/master/01parse_ndl.py
- ^v 例えば、大学を中心に普及している丸善雄松堂「Maruzen eBook Library」は機関向けサービスなので、「一般ユーザーが誰でもアクセス可能な電子書店」とは言えない。
https://kw.maruzen.co.jp/ln/ebl/ebl_01.html
- ^{vi} BOOK☆WALKERヘルプ「取り扱い書籍の一覧はありますか？」
<https://help.bookwalker.jp/faq/301>
- ^{vii} ITmedia eBook USER「『BOOK☆WALKER』を徹底解剖する」（鷹野凌：2013年10月8日）
<https://www.itmedia.co.jp/ebook/articles/1310/08/news007.html>