

デジタル・ネットワーク社会における出版物の利活用の推進に関する懇談会  
技術に関するワーキングチーム（第7回）議事要旨

1 日 時 平成22年6月17日（木）17：00～19：00

2 場 所 総務省第1特別会議室

3 出席者（敬称略）

岩浪 剛太、植村 八潮、宇田川 信生、岡本 明、小川 恵司、尾崎 常道、佐藤 陽一、下川 和男、杉本 重雄、武田 英明、田中 久徳、常世田 良、中村 宏之、新居 眞吾（権正 和博構成員代理）、野口 不二夫（稲井 幸治構成員代理）、萩野 正昭、林 直樹、坂東 浩之、平井 彰司、船本 道子、堀口 宗男、丸山 信人、室田 秀樹、八日市谷 哲生（風間 吉之構成員代理）

4 議事概要

- (1) 松田情報流通振興課統括補佐より、堀口宗男構成員の交代について紹介があった。
- (2) 田中構成員から、資料7-1「国立国会図書館における平成22年度全文テキスト化実証実験の概要」に基づき、説明があった。
  - ・今年度の国立国会図書館における全文テキスト化実証実験について紹介。
  - ・これまで、国立国会図書館と幾つかの出版社で、全文テキスト化の課題検討の必要性についての話があり、その方法について話合ってきた。
  - ・今年度は、昨年度の補正予算を繰り越して大規模な書籍のデジタル化を進めており、効率的にテキストデータを整備していくことを国立国会図書館の課題として、例えば障害者の方への提供も視野に入れて、OCRの処理等技術的課題について検証、取組を計画していたところ。
  - ・それとは別のフェーズで、全文検索に係る課題を検討しようという話が出版社からあった。今年度は、国立国会図書館の実験の一部として、出版社の協力を得ながら、古いものからつくるテキストデータと新しく様々な形で流れていく電子書籍データを一緒に検索できるような実験、課題を検討してみようということで、それを図にしたのが資料7-1としてお示ししたものの。
  - ・国立国会図書館としては、今まで書誌による検索サービスを提供してきたことも踏まえ、書誌から検索して、そこで全文検索ができる場合、どういうことができるようになるかといったことも検討してみたい。

- ・また、従来から国立国会図書館でもやってきている、件名、分類等資料の中身にかかわる知識の組織化を、従来の書誌情報と組み合わせる形で、フルテキスト検索ができるかどうかといった可能性があるのかも検討したい。
- ・出版社からは、様々なジャンル、データを提供していただき、どのように効率的に検索システムを組めるかも検討していく。
- ・検索結果は、いろいろな形での活用が考えられるが、そういったデータベースの連携も含めた検討を今年度はしてみたいと考えている。
- ・来年度以降は、行政の支援等もできればいただきながら、今年度の課題を踏まえた形で、公共的な観点、商業的な利用の可能性も含めて、全文検索が具体化していくようなアプローチを続けていければと考えている。

(3) 資料技7-2「デジタル・ネットワーク社会における出版物の利活用の推進に関する懇談会報告(案)」に基づき、議論を行った。

これについて、以下のやりとりがあった。

- ・図書館と民間の役割分担の中で、「実際に図書館で貸し出されている本はベストセラー本や娯楽本が多く、著作者や出版者、地方の書店などへの経済的な影響は少ないため、電子出版を図書館が貸与(配信)することについては、抑制的であるべきとの意見も出された」との記述があるが、公共図書館での貸し出しがベストセラーや娯楽本に偏っているという事実はなく、事実誤認ではないか。
- ・その部分については、データ等のとり方、解釈の仕方も含めて、意見があり、どう評価するかというのは意見が分かれている。またこの部分に限らず、ほかの論点についても両方の意見を併記している。その意見を踏まえて、今後、例えば図書館の役割について検討会を設けて、論点整理をして、正面から議論していく必要があるという方向性にしている。
- ・技術ワーキングチームでの公共図書館等を代表する意見が、必ずしも利活用ワーキングチームの場に反映される形になっていないという指摘もあるが、そういったことも含めて、改めて図書館の声も入るような形で、今後議論されていくということなので、基本的にはそちらのほうに検討をゆだねていくということではないか。
- ・「全文テキスト検索」という言葉の定義がわかりにくいのではないか。検索して得

たいものは全文テキストではなく、例えば「猫」という言葉を入れただけで『吾輩は猫である』という本が取り出せるかということ。目的があつての検索なので、「猫」といれたときにその書籍が検索できることが重要であるということがわかるように工夫したほうがよい。

- ・海外の出版物に自由にアクセスできるようにするとともに、日本の出版物を世界に発信するために、翻訳まで含めたワークフローの確立についても触れるべきではないか。
- ・「メタデータ」、「書誌情報」の言葉の使い方を整理すべきでは。例えば書誌情報を含むメタデータといった説明の仕方など。
- ・「文字セット」と「文字フォント」の言葉について、「字形をあらわす文字セットとスタイルをあらわすフォントセット」などの補足説明が必要。
- ・DAISY (Digital Accessible Information System) は、子どもの読書と教育への活用だけではなく、障がい者へのアクセシビリティを確保するためのものという前提を記述すべき。技術的な言葉では、「SMIL (Synchronized Multimedia Integration Language) 形式」が適切ではないか。

(3) 資料技7-3 「「デジタル・ネットワーク社会における出版物の利活用の推進に関する懇談会報告(案)」の具体的施策の実現に向けて」に基づき、議論を行った。

これについて、以下のやりとりがあった。

- ・書店を通じた電子出版と紙の出版物のシナジー効果について、携帯端末販売を参考に、書店が電子書籍端末を売っていく仕組みをつくれればよいのではないか。
- ・端末レベルで売ることになると、量販店には勝てない。量販店・書店が同じ値段で勝負できて、かつインカムが入ってくるという仕組みがあれば、普及にも役立つのではないか。
- ・書籍流通で一番問題となっている返品の問題、在庫とも関係してくる問題ではあるが、アメリカのバーンズ・アンド・ノーブル、イギリスのウォーターストーンズがメーカーと組んでいるという実態もある。今後、検討はオープンな形でやっていきたい。

- ・これまで書店と一番協力しながらやってきたのは出版社。出版社と書店が日本の出版文化を支えてきた。デジタルの世界においても、これまでと同様に、出版社と書店が協力して、読者によりよいものを提供できる方法を一緒に考えていきたい。
- ・実際に具体的な施策に動かしていくためのものについては、ここでの議論をキックオフとして進めていければと考えている。今回の報告では、関係者における会議の立ち上げ、取組を行うという形の提言になっているので、そこはまさに民間での取組を進めていくということに関係者でコンセンサスをとってもらいたい。また、行政としても支援していきたい。
- ・国立国会図書館の全文テキスト検索実証実験の賛同出版社は、今後広く呼びかけをする予定。できるだけ様々な種類、雑誌、専門書、学術書も含めて実験を行い、全文検索の有効性、国立国会図書館がサービスをすることが適当かどうかも含めて実験、議論をしていきたい。
- ・外字・異体字の問題について、各団体と相談して進めていかないと最終的な電子書籍は完成しない。
- ・中間フォーマットの委員会とEPUBの委員会は、重複、不可分の部分が相当あり、その調和、調整を行っていく必要があるのではないか。
- ・中間フォーマットを策定、デザインするとき、ターゲットフォーマットを見てつくるのではなく、書籍の電子構造、XMLの構造をどうつくるかを考えてつくればよいのではないか。
- ・あげられた課題について、個別的な問題だけではなく、この委員会のような統括した場が必要なのではないか。少なくともフォーマットについては、全体の中で、EPUB、交換フォーマットがあるという枠組みのほうが議論しやすいのではないか。書誌情報、外字問題等技術的課題については、お互いある程度開示し、競争する場がなければ進まないのではないか。
- ・DRMの運用の方式に関する実証実験もやるべきではないか。
- ・中間フォーマットのライセンスフィーは、非常に重要な部分。ここで示されている交換フォーマットは、現状、出版社で使用されているライセンスのもとにある2つのファイルフォーマットを統一して、その中で公的標準にするという枠組みまで。公的標準にするにあたってのライセンスの形は、各社、各団体の考え方になる。また、ファイルフォーマット周辺におけるビジネスは、それとは別に各自が考えることになる。

- ・第3回親懇談会で、一定の取りまとめが報告（案）の形で出されるが、計画的な施策展開を図る観点から、その後も必要に応じて懇談会及びワーキングチームにおいて進捗状況のフォローアップをしていきたい。

以上