

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1回）議事要旨

1 日 時 平成22年4月15日（木） 15:00～17:00

2 場 所 三田共用会議所第三特別会議室

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、下川 和男、杉本重雄、高橋 靖明、武田 英明、田中久徳、常世田 良、中村 伊知哉、中村 宏之、新居 眞吾、野口 不二夫、萩野 正昭、林 直樹、坂東 浩之、平井 彰司、船本 道子、丸山 信人、室田 秀樹、八日市谷 哲生

4 議事概要

(1) 内藤副大臣による挨拶

○ 内藤総務副大臣より以下のとおり開会の挨拶があった。

- ・ 懇談会を立ち上げた趣旨は、一言で言えば、表現の多様性の確保である。アメリカでは大きな資本を持つところが、縦系列の電子書籍の市場をつくり上げた。私は、資本が多いところも、また、少ないところも、多くのプレーヤーが参画する中で、我が国の出版界・出版市場というものをもっともっと多様なものにしていこう、それでこそ健全な電子出版市場というものが形成されるのではないのかと考えている。
- ・ もう一つは、我が国には、国立国会図書館に膨大な知のインフラがあるが、残念ながら今、地理的な理由もあって、やはり東京近郊等の一部の人しかそれにアクセスできない。ブロードバンドを使って、だれもが皆、簡単に膨大な知のインフラにアクセスする、そのような環境を整えることも国が果たすべき役割ではないのか。そのような思いで、この電子書籍に関する懇談会を立ち上げさせていただいた。
- ・ 皆様方には特に技術面での議論をしていただくわけだが、本当に多くのプレーヤーが参画できるようにし、国会図書館と我が国が持っている公文書も、この1つの共通の基盤の中でだれもが皆容易にアクセスできる、そんな環境をつくり上げていくために、これから4月、5月、6月と、決して長い時間はないが、3カ月という短い時間、どうか我が国の将来の出版市場というものを健全なものに育て上げるためにも、皆様方の忌憚のないご議論を心からお願いを申し上げて、政府を代表してのあいさつ、お願いに、そしてまた御礼にかえさせていただきたい

(2) 「検討テーマ（案）」についての説明

○ 安藤情報流通振興課長から資料技1-2「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 技術に関するワーキングチーム」検討テーマ（案）」に基づき、以下のとおり説明があった。

- ・ 資料技1-3のとおり、検討事項を2つに分類している。一つは1のところにあるビジネス流通環境に関する技術、もう一つが4ページ目にある、公共的なアーカ

イブの構築である。

- 最初のビジネス流通環境について、先般の親懇談会での議論等々から、「オープン型の電子書籍環境」の実現に必要な技術的課題を検討していくべきではないか。「オープン型電子書籍環境」は今回初めて出しているものであり、厳密に定義づけることはなかなか難しいが、できるだけ多様なプレーヤーが参加できるようなもの、あるいは利用者が国内外の豊富なコンテンツに簡便、それからできるだけ自由にアクセスするようなこと、そういった状態を表現して「オープン型電子書籍環境」と言っているが、そういったことを実現するという点で検討を進めていくべきなのではないか。
- 検討の方法ということで、書籍の種類に応じた検討が必要ではないかと考えている。書籍と言っても非常に多くの種類のものを含んでいる。その書籍の種類によって、そもそも制作の過程も非常に異なり、また、利用の仕方等々も異なってくるので、書籍の区分例ということで5つほど書いているが、どの書籍を念頭に置きながら議論していくのかといったことを常に意識しつつ議論をしていく必要があるのではないか。
- 端末の種類に応じた検討ということがある。電子書籍の端末に関しても、さまざまな種類がある。また、端末の種類に関連づいて、配信形態や課金方法、すなわちサービスのプラットフォームや流通形態が異なる。端末自体の機能や性能、プラットフォームの機能や性能、流通形態に関しても十分に念頭に置きながら議論を進めていく必要があるのではないか。
- 技術的項目に関しては3ページ目に、参加者からいただいた意見をもとに、事務局のほうで整理をしている。この事項に限定するという趣旨でもなく、議論の方向性を示しているということではないということをご理解をいただきたい。

○ 構成員より、以下の質問があった。

- マルチメディア書籍について、ワークフローについて触れられているが、でき上がったものがどうであるということではないと考えて良いのか。

○ 事務局から以下の回答を行った。

- つくり方だけではなくて、でき上がったマルチメディア書籍も入っている。

(3) 下川構成員から資料技1-4「電子出版フォーマットとEPUB、JepaX」に基づき、説明があった。

- 電子出版フォーマットは数多くあり、大きく制作会社、出版社、配信会社と分けて、素材用のフォーマット、原稿用のフォーマット、交換用というものを作成している。蓄積交換用で持ち歩くもの、EPUB等、配信して実際に見るもの、読者を管理するものというようにいくつかの用途に分かれると考えている。
- こういったフォーマットのうち、日本電子出版協会が昨年11月から、策定を始めたものがEPUBというもの。電子出版協会のプロジェクトのミッションは、欧米でデファクトとなりつつあるEPUBの日本での普及啓発と、日本語組版を早

期に実装してもらおうということであり、EPUBは、欧米でさまざまな会社に採用されて標準になりつつあるが、日本語の組版について、ほとんど議論がされていない。

- 現在のEPUB 2.0は、さまざまな標準を一緒に入れたような形式になっていて、電子出版協会が要求書を提示したのは、XML 1.1とCSSのレベル2に対する14項目。
- 1999年にEPUB 1.0が策定されて、2002年に1.2になったが、IDPFという団体が仕様を策定したという経緯があり、2007年に今回の2.0が策定され、それをアドビ、ソニー、バンズ・アンド・ノーブル、グーグル、アメリカの図書館が採用した。そして、最近アップルがiPadで採用した。
- 3-2の資料の中にHTML5とCSS3と書いてあるが、このCSS3は縦書きで段組を行い、詳細な印刷組版に近いことを画面組版でやろうとしているもの。そのうち一部の機能を先取りして、縦組みとか右開き、柱、ノンブル、キャプション等を今回のEPUBの日本語版の要求仕様の中に入れた。
- EPUBは、とてもリッチなフォーマットと思われやすいが、基本はリフローという技術を中心に策定されたものであり、文字サイズを変える、もしくは画面サイズの変更によってリフローされる、再描画されるということが特徴であり、自分の読みたい文字サイズで本が読めるという機能を一番重視しているもの。したがって、雑誌のきれいなレイアウトを表現したり、写真は必ずこの位置にないといけなとか、そういうことはできない。
- 今回、日本から要求仕様を提示したところ、中国、韓国についても、例えば、外字についてどうするのか等、いろいろな問い合わせが来ているため、日中韓の3カ国で会合を開催することになり、連休中にソウルで会合を行う。今後も会合を重ねていって、日中韓で合意をとった後、国際標準化も目指したいと考えている。
- EPUB以外に電子出版協会では、JepaXという電子出版協会が1999年に策定したものがある。策定から11年たったが、交換用の中間フォーマットなので、これから活きるのではないか。日本印刷技術協会、JAGATでも、2010年の4月に、JepaXのドキュメント作成ツールとJepaXの自動組版ツールをオープンソースで公開している。

(4) 萩野構成員から、資料技1-5「designing for Webkit?」に基づき、説明があった。

- 1つの本のデータというのは、まず、読者のさまざまなデバイスで閲覧されなければならない。2番目として、いついかなるときや場所においても閲覧は保証されなければならない。それから3番目として、いわゆる貸出と販売が備わっていないなければならない。それから4番目として、出版社の意思が尊重されなければならない。
- 電子出版の歴史というのはとにかくデバイスの歴史だった。そのデバイスに合わせていろいろなフォーマットができ上がってきた。このために、そのデバイスの撤退と同時に、関係するあらゆるソフトウェア、コンテンツがひっくり返るという歴

史がずっとあった。さまざまなデバイスで閲覧されなければならないというのは、これは読者の立場に立ったときに非常に重要な観点。

- Lending と Vending について、現在の本が置かれている環境は、まさにこれである。原則としては貸出と販売というものがついていなければいけない。
- 出版社の意思が尊重されなければならないという点について、価格の問題等がすべてコントロールされてしまうことやこんな表現があったら出ていけというような流通の構造であってはならない。
- そう考えて、ドットブックというフォーマットをつくってやってきた。今、現状でできることをお話しすると、PC上でこれをストリーミングで見るというようなことや iPhone でもものを見るということ等が可能である。
- ボイジャーでは T-Time というビューアを持っている。このビューアで見ると、縦書き、字の拡大、横書き、いろいろなことができるようになっていく。先ほどあらゆるデバイスでものが見えなければいけないと言ったが、この T-Time というビューアを使って、例えば iPhone のサイズ、あるいは携帯のサイズに合わせて表示をして、それを書き出して、デバイスで見るというようなことが可能になっている。ただ、これは1つの画像的な表示でしかないので、検索等が追従できない。
- 紙の本というのは、視覚障害者にとっては本ではないという1つの現実がある。電子的な形をとることによって、視覚障害者に対する読み上げが非常に容易になってくる。例えば、EPUB等では、いわゆるデージーというフォーマットを使用するという方法をとっているが、ドットブックは、データを視覚障害者が使われている一般的な読み上げエンジンに渡して読み上げるという方法をとっている。これがなぜできなかったかという点、ドットブックというのにはDRMがかかっているから。
- 現状のDRMは読者にとって1つもいいことがない。また、紙の本を出している以上は、DRMをかけたからといって、このことによって有効に作用しているかという点、何もしていない。DRMがかかっている点、読み上げができないという問題があった。
- EPUBというのが世界的な基準となるという方向が出てきているので、ドットブックというのは、EPUBへの変換というのは非常に大きな問題になっているが、そもそもドットブックは、XHTMLベースのEPUBへの変換は非常にスムーズに行く。
- ただ、日本語特有の幾つかの問題がある。縦書き、ルビ、縦中横、禁則、圏点、外字表示等、非常に我々にとって大事だが、なかなか世界の標準というものに入れるということは難しい。IDPFというEPUBの策定委員会で理解を得るのがなかなか難しいという問題がある。
- ボイジャーは、アメリカのインターネットアーカイブと共同して動いているが、Open Publication Distribution System というものをオープンライブラリーという形で持って、これを基本的にパブリッシャー、オーサー、ライツホルダー、ブックストア等に対して提供していこうとしている。これは、あらゆるデバイスに自分

たちの作った、出版する者が作ったデータを配信していく1つの未来図。

(5) 尾崎構成員から、資料技1-6「(「放送による新聞・雑誌等のデジタル配信 All Media In One(AMIO)プロジェクト」について(概要)」に基づき、説明があった。

- AMIOフォーラムは、すべてのメディアが1つにつながるというような思いを込めてつくられたプロジェクトで、デジタル放送を利用した新聞・雑誌の紙メディアの完全デジタル配信を目指し、そこにかかわる技術的な課題、制度的な課題などについて検討を進めていくことを2つの大きな検討テーマとしている。
- 放送の特徴というのは、特定の時間帯に集中して配信し、廉価に大多数に一時的にコンテンツを配信するのに適しているということ。基本的なコンセプトは、理想形としては、つくり手側からすれば1つのワンソースで、たくさんのデバイスに対して自動的に配信されること。ユーザーから見れば、どんなデバイスが手元にあったとしても、すべてのコンテンツにアクセスできる状態をつくること。
- 一方で、現状はいろいろなフォーマットがあり、いろいろなデバイスがある。その中で、例えばポータルが、デバイスにくくりつけられてしまうという、いわゆる垂直統合ができつつある。これを何とか今の時点で回避しなければいけないというところがこのワーキングチームのポイントと考えている。
- AMIOプロジェクトで構築したシステムがまさにそれであり、つくり手側は、なるべく今のワークフローを変えないように入稿する。デバイスは切磋琢磨してどんどん出てくれば、マーケットもそれを評価すればいいのではないか。ただ、フォーマットもいろいろあるので、その「差分」を流通側で吸収してしまおうというのが、今回のAMIOプロジェクトの一番重要なコンセプトと思っている。
- 入稿の段階で、今のワークフローを極力変えずに、一手間かけてもらえれば、変換がしやすくなるので、その方法をいろいろと考えた。例えば、InDesignソフトで構造化を行えば、テキストや画像等の各要素がどのように配信されるか、各要素の並び順をどうするか、写真、キャプションのグループ化をするか等、一手間だけかけて、非常にスムーズに行くのではないか。
- そのために、コンテンツ制作ガイドライン、あるいはプレゼンテーションガイドラインという形をまとめるという活動を行った。
- また、コンテンツメタデータを付与した。例えば、座標、記事エリア情報の指定を行うようなメタデータ、オーバーレイ表示する場合のテキストの透明度の指定を行うようなメタデータ、そういうメタデータを付与した。
- また、配信メタデータの充実により表現力を向上させることができる。記事エリアで自動フィットし、透明レイヤーをつかって文字を重ねてオーバーレイ表示等もする。縦書き、横書きの問題もそこで解消できる。
- 1つのワンソースで、端末市場はとにかく切磋琢磨できる状態にする。その間に立った流通過程で、「差分」の吸収を行い、コンテンツを配信し、最終的にはユーザーの選択の自由度、アクセスの自由性を高めたと実験をしてきたというのがこのAMIOフォーラムの活動。

- (6) 田中構成員から、資料技1-7「デジタル・アーカイブの標準化に向けて」に基づき、説明があった。
- 公共系のデジタル・アーカイブの国会図書館でのこれまでの取組ということで、標準化に向けての今までの取組について簡単にご紹介させていただく。大きくは、既存の出版物のデジタル画像をつくるということと、インターネット等で流通しているウェブサイトやオンライン資料をアーカイブしていくことの二つがある。
 - デジタル・アーカイブのシステムとしては、電子書庫というような形のストレージシステムに永続的識別を付与して、基本的に集めたコンテンツは永続的に永久にわかる形で識別される形で残していくというのが基本。
 - 標準化に関する考え方としては、2004年に電子図書館の中期計画をまとめているが、その中で関係機関、デジタル・アーカイブを進めていく国内外の機関との連携ができるように、データの共有や相互運用性を確保する目的でメタデータや長期的な保存にかかる技術標準に積極的にかかわっていくことを決めている。
 - さらに、ポータルによるサービスというのを1つの大きな柱にしているので、メタデータを共通のものにして、ポータルを構築するというのもう一つの柱になっている。
 - データの標準化の一つはメタデータだが、従来から書誌データのMARCの標準化を進めてきた。今、国際的な交換ということを視野に置いて、JAPAN-MARCのフォーマットをMARC21という、アメリカの議会図書館が策定したフォーマットに変更することを2012年に進める予定。
 - 文字コードもJISコードからUnicodeに変え、電子的なもののメタデータとしては、ダブリンコアに準拠したDC-NDLという、国立国会図書館メタデータ記述要素というものを策定した。現在は、2007年に改定したバージョンでやっているが、現在、さらにまた改定作業を進めているところ。
 - 既存の紙からの資料をデジタル化するという点については、手順等についてなるべく標準化を果たすということでデジタル化の手引きを作成している。
 - 画像については、一応、保存用の画像、提供用の画像ということで、それぞれ標準を定め、さらにオプションを定めるという形で、解像度等について基準を定めている。
 - 画像から全文テキスト化できるような、テキストデータをどうやってつくっていくかという問題があり、現状、古い資料にかかるようなOCRが十分でないので、試行錯誤の形で、試行実験を幾つかやっている。
 - さらに、作成したテキストデータについても、プレーンのテキストのままなのか、それともタグづけして構造化していくかということもこれから検討していかなければいけない。
 - データの流通にかかわるプロトコルの標準について、横断検索や統合検索ができるようにするために、メタデータをハブとしてくるプロトコル、横断検索用のプロトコルを全部実装した形で、さらに機械的な連携ができるように、今後はシステムのAPIを公開していくというのが標準になっていくという形になると思う。

- ・ 今、納本制度の審議会で、商用的に流通しているオンライン資料の納本制度としての収集についての検討を始めている。電子書籍の流通フォーマットというのは現状でもいろいろなものがあるが、版面デザインを保持した形で集めることと、内容を長期的に保存していくということが求められる。しかし、その2つはなかなか両立しがたい部分があるというのが現状であり、大きな課題と認識している。
- ・ 公共的観点では、アクセシビリティを確保していくということで、読み上げ等に対応できるようなフォーマットは何らかの形で維持していく必要がある。その場合に、DRMを解除することやフォーマット変換等を含めて同一性保持等にかかるような権利上の課題が、これからのもう一つの問題としてあるだろう。
- ・ 総務省の協力を得て、いわゆる限定的な利用のための配信の実験等を行ってきている。

(7) 自由討議

- ・ 図書館も、国立国会図書館と公共図書館では、役割が違う。また、国立情報学研究所は、学術情報に限定されるので、立ち位置が違う。長期にわたる保存といった話は国立国会図書館や国立公文書館では非常にシリアス・重要な問題として位置づけられているのではないか。そういう意味で、内容の保存なのか、もともとの出版物そのままの形で100年もたせるのかといったような意識は、国立国会図書館・公文書館と、公共図書館でかなりスタンスが違うのではないか。
- ・ EPUBの次のバージョンについて、その策定を待たずに、例えば中国でだけ仕様を実装したイーブックリーダーが先行して発売されるというような可能性もある。EPUBの次のバージョンの策定を待たずに、電子出版協会がIDPFに要求しているEPUBの仕様に即してベンダがビューアを発表するという作戦もある。
- ・ 中国は、中国の漢字の歴史を電子化するときに縦書きを捨てて、横書きのみにした。これは文化的な意味で大決断。世界中が横書きのみになっているが、まず、日本は、文化の継承として、電子化しても縦書きは捨てないということを確認しておかないといけない。
- ・ EPUBというのは大変緩やかなルールなので、EPUBの上にさらにガイドラインを提供しないと、専用端末などできちんと読めないというようなこともある。対象とするデバイス、拡張性を確認した上で、縦書き、ルビ、横書きをどのように考えるか、大方針をきちんと確認をした上で、進めていく必要がある。
- ・ 縦書きは当たり前。今、活字系電子出版のほとんどが携帯メインになっているが、携帯という小さな画面と小さなリソースの中で可能なもののみを提供しているというのが実態。シャープ、ボイジャーが日本における実際のデファクトになっているが、出版社はそれを使い分けている。
- ・ これからは紙の本と電子書籍同時リリースという時代がやってくる。同時リリースをする際、紙の本のデータを何らかのフォーマットに変換できなければ、同時リリースは大変難しい。EPUBは、日本のいわゆる文芸出版社、書籍出版社に対して事前の相談を全くしていないが、あたかもデファクトになったというような言い方で、進められてしまうというところに危惧がある。

- 今まで本を作ってきた印刷会社と出版社主導で、著者を交えて、日本なりの必要なフォーマット、あるべきフォーマットというのをまずは考えて、その理想形にどうやって近づけていけるのかというようなアプローチをとることも必要ではないか。
- 実際、今、日本の電子書籍は携帯がほとんどで、コミックが主体で、アメリカの市場より大きい。今後はキンドル・iPad的なビジネスがやってきて、おそらくあつと言う間にそちらの流儀になっていくのだろう。
- 技術検討する際に留意しておかなければいけないのは、日本のデジタルコンテンツ配信ビジネスは総じてユーザーに対してコンテンツを販売していないケースが多いということ。電子書籍の最近のアンケートを見ても、コレクションできないとか、買った気がしないという回答が上位に必ずあるはず。権利処理あるいはDRM等の問題があり、電子書籍を売っているのか売っていないのかわからない状態。この状態だと、キンドル、あるいはアップル型が上陸してくるとひとたまりもない。やはりそれを実現可能とする技術体系を念頭に置いて進めていただきたい。

(以上)