

「電子出版インデックス情報データベース」

コンセプトガイドライン

2011年5月版

社団法人日本雑誌協会

1 目的・位置づけ

電子出版のインデックス情報及びインデックス情報データベースの基本的なコンセプトの指針を示すことを目的とする。

読者（国民）の電子出版の検索を容易にし、出版社（者）がデジタルの新しい企画・編集ビジネスを課題解決のための参考点を示し、配信流通事業者等が電子出版と連携した多様なビジネスへの展開のための参考点を示すことを目的としている。

2 背景と課題

我が国の出版市場は、書籍・雑誌の売上は、雑誌は13年連続してマイナス成長であり、書籍は3年連続してマイナス成長である。2009年の書籍・雑誌の推定販売金額は21年ぶりに2兆円を割り、1億9356億円となり1996年に過去最高の2兆6,500億円まで伸びた売上が連続して減少している。近年の出版社（者）数減少や書店数の減少に見られるような出版流通の変化、国民の活字離れ、電子メディアの成長による情報手段の多様化などが要因として考えられる。

一方、電子出版市場における取り組みは極めて活性化している。我が国においても新たなプラットフォーム向け電子出版市場の拡大により、2014年度には2009年度比約2.3倍の1,300億円程度の市場規模と予測されている。電子出版が持つ技術的な可能性を活かすことで、より多くのコンテンツ（出版社（者）による新たなコンテンツの創造など）を、より多くのサービス（配信流通・端末会社などによる新しいビジネスモデルの創造など）で、より多くのサービス利用者（読者）を獲得していくというように、電子出版市場の拡大が期待されている。

その中で、電子出版が持つ技術的な可能性の一つとして、紙の出版物では想定できなかった目次・記事単位で細分化されたコンテンツ（以下、「マイクロコンテンツ」という）を流通させ、利用者がニーズに応じて読みたいマイクロコンテンツを閲覧・参照する環境を構築することが可能と考えられる。

現在、先行事例である学術分野においては、国際学術雑誌の大半は電子化され、記事単位に細分化されたコンテンツの電子配信が行われており、検索容易性・本文到達性の向上といった顕著なメリットが獲得でき、既に一般化している。

しかしながら、一般の電子出版の分野においては、国民が読みたい記事を選択して読めるなどの「オープン型電子出版環境」整備の研究が不十分であり、マイクロコンテンツ単位でコード（ID）を付与して、その促進を図る仕組みなどについて検討する必要がある。

これらを鑑み、オープン型電子出版環境を実現するためには、以下 3 項目の整備が必要と考えられる。

- (1) 多くの読者（国民）に、電子出版を選択して読める環境を構築するために、電子出版の流通管理を円滑化するための、目次・記事（マイクロコンテンツ）単位のコード体系の整備。
- (2) そのコード体系に基づいて、読者（国民）に、電子出版の検索を容易にし、出版社（者）がデジタルの新しい再編集ビジネスを実現するための「電子出版インデックス情報データベース」仕様の策定。
- (3) そのインデックス情報を利活用し、多くの出版社（者）ビジネス参入障壁を下げ、配信流通会社などとの流通を促進し、読者（国民）に、多くの電子出版コンテンツにて提供する仕組みの検討。

本ガイドラインでは、上記（2）の課題について述べる。

3 課題解決のための方向性（仮説）

現状ではマイクロコンテンツを検索・再編集する仕組みがまだ不十分である。特に、検索においては、コンテンツの全文検索だけであり、読者（国民）は、出版物の内容を把握できないケースが多く、別の検索方法が求められている。目次・記事単位の流通というのは、雑誌や書籍が持つパッケージの価値を失う可能性もあるため、記事自体に何が書いてあるのか認識するための情報となるメタ情報の存在が非常に重要になる。メタ情報を利用し、記事と読者の効率的なマッチングを行う仕組みが本課題解決への重要な要素になりうると考えている。

また、電子出版の提供方法については、現在の紙と出版物と同様のパッケージ単位に加えて、新たなサービスとして、目次や記事単位でのサービス提供が可能と考えられる。ICT の利活用により、目次や記事単位で管理することにより、異なる出版社（者）間の記事を再編集した電子出版物や、過去の記事から共通したジャンルやテーマ（例：エコや富士山といったテーマ単位、コラムといったコーナー単位など）を組み合わせた再編集された電子出版物などの制作及び提供を、新たなサービスとして創造し、従来の紙の出版物では制作及び提供が困難であったコンテンツを組み合わせた新企画・再編集ビジネスも活性化させ、電子出版市場をゼロサム（zero-sum 利得の合計がゼロ）からプラスサム（plus-sum 利得の合計がプラス）の市場にする可能性を有している。

(1) 記事の構成要素に対する検索機能の整備

出版社の編集者が、電子出版等の企画・編集を行う際に生じる検索要件として、以下のような点が考えられる。

- ・タイトル、大見出し、キャプションなどの構成要素に対して検索したい
- ・構成要素に対して重み付けを指定して検索したい
- ・検索された記事は発行年月が新しいものについて優先度を高くしたい

そこで、上記要件を解決する仮説として、記事に対して構造解析を施すことで記事の構成要素を抽出し、検索の際に利用することが挙げられる。検索の際に構成要素や重要度の指定が可能となり、記事構成を意

識した詳細な検索が可能になると考えられる。

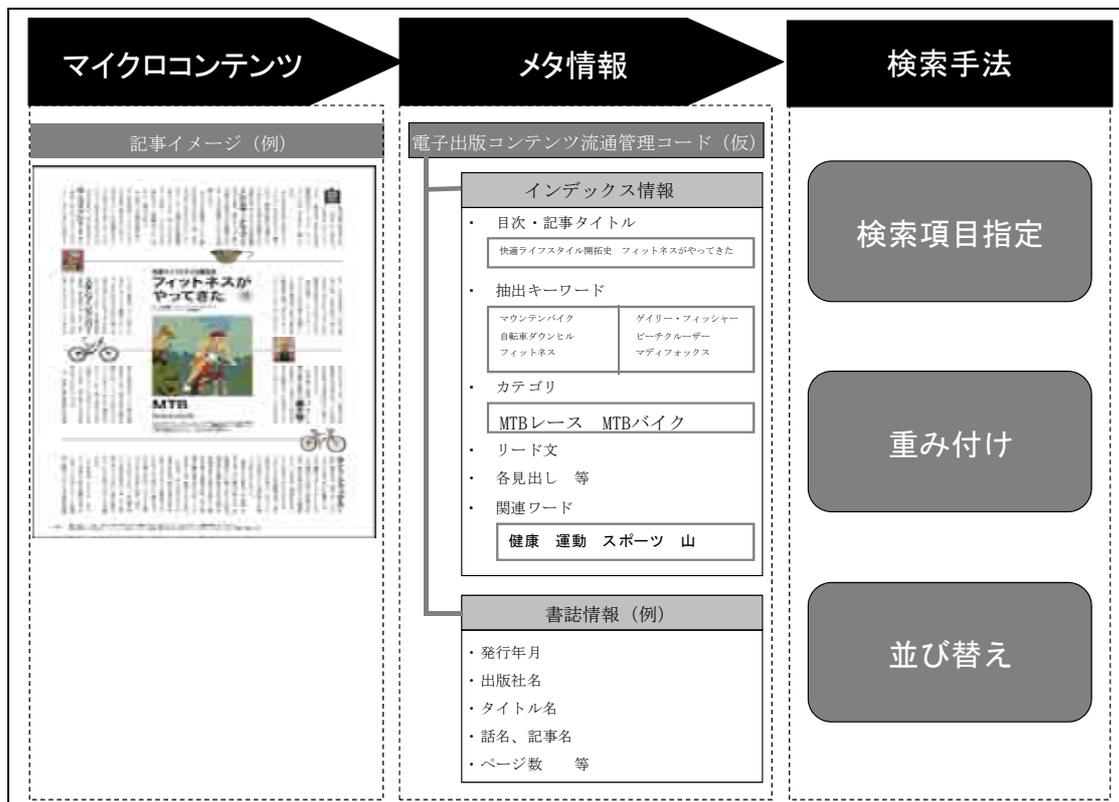
(2) 電子出版コンテンツ流通管理コード（仮）を包含した書誌情報との連動

マイクロコンテンツ単位で付番される電子出版コンテンツ流通管理コード（仮）を包含する書誌情報と構造解析した情報が連動することにより、メタ情報として発行年月などの書誌情報を活用し、検索結果へ反映することが可能になると考えられる。

(3) 記事の持つ意味を加味した検索機能の整備

目次・記事単位という狭小範囲の単語に対する全文検索では、効率的なマッチングが望めないと考えられる。探したい記事が指定した単語を含んだ記事でなければならず、事前に記事の中に含まれる単語が分かっている状態であるならば良いが、そのようなケースは稀である。また、文字の少ない記事であった場合には更に必要である。そこで、記事の持つ意味を示唆する単語による検索機能（インデックス情報検索）を行うことで、探したい記事への到達性向上が期待できる。

この場合、記事の中に指定した単語を含んでいる必要はなく、記事が持つ意味にマッチングすれば検索結果とすることが可能になると考えられる。



資料1 構造解析及びインデックス情報検索イメージ

このような検索の意図をコンピューターに理解させ、要望に応じた結果を導き出す技術は、セマンティック技術と呼ばれている。これは、今までの検索手法で主流だった、指定した単語に対するマッチング技術とは異なり、指定した単語から本来要望している事柄を理解して検索結果を導き出す技術である。

次に、例を用いて説明する。

例えば、Aさんが友人へ贈り物をしようと考えている。友人は甘いものが好きなのでスイーツに関する記事を探したい。この場合は、検索ワードとしては「スイーツ」を指定し、「スイーツ」という単語を含む記事が検索結果として妥当と見なされる。しかし、Aさんの意図を理解すると「甘いもの」に関する記事を探しているので、「パティシエ」や「デザート」など甘いものを示唆する単語を含む記事も検索結果として欲しいものである。このように単語に対して関連性の深い単語を類推して検索結果に対して補完する技術である。

上述した仮説を包括的に実現するものとして、インデックス情報データベースのコンセプトを作成した。

4 電子出版インデックス情報データベースの必要性

インデックス情報データベースの実現により、各プレイヤーに以下の様なメリットが生まれる。

(1) 出版社（者）

- ・新しい価値あるコンテンツを提供する新企画やデジタルコンテンツを組み合わせた再編集など、電子出版市場において新しいビジネスの拡大が可能となる。
- ・パッケージ単位だけではなく、マイクロコンテンツ単位での提供が可能となることにより、国内市場だけではなく、多様なニーズを求める海外市場に対しても、電子出版の提供が拡がり、海外向けの電子出版サービスへの展開が可能となる。

(2) 配信流通事業者

- ・規模を選ばず様々な配信方法によるデジタルビジネスが可能となる。また、標準化されたオープンなAPIを使い、リアル書店を含むどの企業でも独自の付加価値を付けてビジネスに参入する可能性が広がる。
- ・パッケージ単位だけではなく、マイクロコンテンツ単位での提供が可能となることにより、その記事に沿った広告をリンクして提供可能となり、多様な電子出版の広告ビジネスが可能となる。

上記の環境を構築することで、国民生活の利便性が向上する。

(3) 読者（国民）

- ・複数雑誌の読みたい記事を選択して閲覧することが可能になり、より個人の属性に沿ったパーソナライズされた電子書庫を創ることが可能になる。
- ・関連コンテンツのレコメンドにより、今まで興味を持っていなかったコンテンツへ興味が沸くなどの相乗効果と、新しい文化や出版との発見の場が増える。
- ・出版社（者）を横断したマイクロコンテンツを組合せた電子出版の購読など、新たな価値をもったサービスを享受できる。

以上のとおり、新しいビジネス及びサービスの創造が期待でき、電子出版市場の全体的な拡大につながると考えている。

5 電子出版インデックス情報データベースの管理運用の留意点

前述の通り、インデックス情報データベースは、より多くの雑誌などが格納され、特定の出版社(者)に限定されないオープンな環境であることで相乗効果を生むものと考えられる。そこで、インデックス情報データベース導入時の管理運用については、以下のような点に留意していく必要がある。

(1) 参加者の作業過多とならない運用

参入障壁を下げ、多くの出版社(者)による参入を促進するためにも参加者の作業過多に繋がる複雑な運用管理方法を避ける必要がある。

(2) コンテンツに対するモラルの維持

多くの出版社(者)による参入が望ましいが、参入に当たり一定レベルの審査は必要になると考える。コンテンツ登録者の特定や運用に際してのモラルの維持などに有効となる可能性が高い。仮に悪意のあるコンテンツに起因して他のコンテンツに悪影響を及ぼし刑事事件に発展するようなケースにも配慮が必要となる。これは、規約や約款などによる制約により有効になると考えられ、これらを維持管理していくための出版社(者)または第三者機関による運用が不可避である。

(3) 一定レベルの監視

オープンな環境ではあるが、一方で利活用方法に対する監視も必要になると考えられる。例えば、違法薬物や条例に抵触するコンテンツなどに利活用されることで、著しく出版社(者)のイメージを損なう恐れがある。そのため、規約の作成や監視などによる運用が不可避になると考えられ、電子出版業界の健全な発展のためにも一定レベルの監視機関の検討が望まれる。

(4) インデックス情報の保護

格納されたインデックス情報に対するアクセス制限の管理を適切に行うことが必要となる。出版社(者)の許諾なく再編集や引用などが行えない運用やしくみ、提供範囲の特定(誌面やジャンル、または期間、緊急配信・停止など)など一定レベルのアクセス制限への配慮が重要になると考えられる。