

## 総論：デジタル化時代の新聞切り抜き業務とは

後藤 嘉宏\*

デジタル化時代の新聞では、記事のデータベース利用が増え、また速報記事が増すことで、記事が価値から離れ、事実の報道が尊ばれる。データベース利用される記事は、新聞紙面上の記事と異なり、何面に記事があるか、見出しの大きさがどの位あるかといった価値づけから解放される。その意味で検索者が主体的に個々の記事に臨める。しかしこれらの記事は、記事の置かれた背景や文脈を失う。各種データベースが充実しているデジタル化時代においても、新聞切り抜き業務が存在している理由も、このような記事の置かれた文脈・背景を重視し、さらにそれらを切り抜きのイメージ情報によって思い出す点に関係しよう。

キーワード：新聞切り抜き、記事データベース、価値、文脈、イメージ情報

### 1. 本論攷の目的

われわれの身の回りの情報環境はデジタル化され、昨今、新聞記事のデータベース利用が盛んになっている。その反面新聞の切り抜き業務も、新聞社の調査部や図書館、資料室の現場では続いている。その理由を探ることで、デジタル化時代の情報の可能性と限界に論及したい。

### 2. 新聞記事のデータベース化への前史

まず、新聞のデータベース化の前史を振り返り、新聞切り抜きとデジタル化された記事のデータベースとの関連性を考える手がかりを掴みたい。

新聞資料は、貴重な情報源である。しかし原紙をそのまま保存すると、スペース的に膨大になる。しかも同じ新聞社のものでも、各本社で記事に相違があるし、同じ本社でも版による差もあり、原紙を網羅的に保管することは不可能に近い。そこで図書館において新聞資料は新聞切り抜きや縮刷版として、さらにマイクロフィッシュ化されて保存され、また最近ではオンラインあるいはオフラインのデータベースを利用することで、現物の保存の代替手段とされ得る。

そこで本稿では特にデータベースに焦点を当てたい。

新聞記事のデータベースは、商用を目的とする前に、記者たちの執筆、編集支援のために作られた。

しかもその前史は新聞切り抜きから始まっている。例えば朝日新聞社は1911年6月索引部を設立し、部長、杉村楚人冠の下、新聞切り抜き業務を開始している。索引部は同年11月編輯局制の採用に伴って調査部と改称された。調査部は95年3月ニューメディア本部と統合し、電子電波メディア局のデータベースセクションとなったが、現在でも切り抜き業務を継続している。他方他の多くの新聞社では近年、切り抜き業務を廃止している。

また縮刷版も調査部長の楚人冠が社内資料として保存される新聞の山の解消のために考えつき、1919年8月の「東京朝日新聞」7月号が最初である。

例えば朝日新聞社においても、現在縮刷版の編集は出版局で行われ、別個の組織が担当するにせよ、流れとしては調査部の切り抜き業務を引き継ぐ形で、縮刷版や記事のデータベースサービスは発展してきた。

つまり記事作りに際しての参照をそもその目的にして切り抜き、縮刷版、記事データベースは作られた。歴史的には、それらは同じ流れのなかにある。

しかし、ではどうして朝日新聞社などは切り抜き業務を依然続けているのであろうか。あるいは図書館でも切り抜きが依然なされるのはなぜであろうか。

『びぶろす』の1950年6月号に新聞切り抜き特集がある。そこで東京大学新聞研究所の初代所長、小野秀雄は、新聞を資料として利用するに際し、「『賢明なる判読』を経なければならぬ」と述べ、報道する側の姿勢を併せ見る必要性を強調する<sup>1)</sup>。他方小野と共に日本の新聞学の草分け的存在であった小山栄三は同じ特集号において、外国の新聞切り抜き図書館の事例を紹介して、大抵の新聞資料の利用者は、報道姿勢に関する

\* ごとう よしひろ 筑波大学図書館情報学系  
〒305-8550 茨城県つくば市春日1-2  
Tel. 0298-59-1322 (原稿受領 2003.8.20)

情報ではなく、事実に関する情報を求めていると指摘する<sup>2)</sup>。

当然編集支援のための切り抜きという新聞社内部の切り抜き業務は、小野のいうような報道する側の姿勢を見ることまで求める。しかし図書館での一般公衆の切り抜き利用は、小山のいうように報道姿勢より単なる事実を求める場合も多い。

例えば1984年12月23日の朝日新聞は、国立国会図書館が人員不足と、「索引のついた新聞縮刷版があり、記事のデータベース化も迫っている」ことを理由に、新聞切り抜き閲覧室を廃止すると報じている。これに国会議員や文化人の反撥があるとも、この記事は紹介している。同年12月26日の朝日新聞では文化人らが「新聞切り抜き継続を申し入れ」たとする記事が載っている。そして12月27日の夕刊では「新聞切り抜きは継続」という記事が載り、外部委託するとされる。85年2月以降、調査及び立法考査局で別に行っていた切り抜きを複写する形で、継続するが、専従の要員もいなくなり、分類法も変わり、それ以前との一貫性を欠くものとなった。そこでさしたる反対もなく93年3月これも打ち切りとなった。「最初に記事が出た際に、その記事が今後どのように発展するか分からないので、どの主題項目に入れるか迷い、その結果、作業が遅延する。また開始当初の1948年の新聞は、ページ数も少なく、対象紙も5紙であったが、51年に対象紙が10紙に増え、ページ数も増加した。そのため、59年には11ヶ月の遅れを見ることとなった」(2003年8月19日、22日の電話聞き取り)。このように国立国会図書館では中止のやむなきに至ったが、各種図書館、資料室、また新聞社で切り抜きが継続している理由はどこにあるのであろうか。

切り抜きの必要な理由として、当然コンピュータ・リテラシーのある者しか、データベース利用ができない、そして高齢者はキーボード操作が苦手であるという問題が挙げられる。しかしそれだけではない。新聞社や企業の資料室のように、専門のデータベース要員のいる組織では、操作が不得手な記者や研究員でも、打ち出した形での情報提供を求め得るからである。にも拘わらず依然切り抜きが行われるのはなぜであろうか。

つまりこの問題の考察は、同じ歴史的な流れのなかに位置する切り抜きと記事データベースとの相違点への考察を促す。

また記者の執筆・編集支援というデータベースの前史に存在した目的と、広く公衆にデータベース化された記事の利用機会を供するという目的にも、大きな相違があり、当然それらの位置づけの究明も、切り抜き

と記事データベースとの相違への考察をさらに促す。

4. でこれらの相違の考察を深めたいが、その前にまず3. で、4. での議論の前提を知るために、新聞記事のデータベースの沿革と多様な姿を概観したい。

### 3. 新聞記事のデータベースの沿革と諸相

#### 3.1 新聞社のデータベースサービスの沿革

新聞社のデータベースサービスは日本経済新聞社の経済情報のサービスから始まっている。それは1970年に開始された日経 NEEDS である。これは記事ではなく磁気テープによる企業財務データを、オフラインでサービスするもので、主な顧客は金融機関、官庁、大学であった。またオンラインによるデータベースサービスは、日経が企業財務情報を扱った72年の NEEDS-TS が最初である。

また記事データベースの商用化は78年5月に日経の NEEDS-IR がオフラインで開始されたのが最初である。オンラインのサービスは日経の NEEDS-IR オンラインが最初で、81年である。朝日新聞のデータベースの構築は84年であるが、商用化は86年4月に行われた。読売新聞はデータベースの構築が86年で、商用化は87年4月から行われた。毎日新聞は87年の構築で、商用化は同年に行われた。

#### 3.2 その他のデータベース

日本新聞協会の「報道各社における調査・資料部業務の現状」調査(1994年)には、切り抜きを廃止した新聞社ではデータベースがテキスト形式であるため、イメージ情報が分からないとの不満の声が紹介されている。また回答のあった80社のうち14社で写真データベースを作り、19社で計画中であった。

このように新聞各社は、検索の便宜からテキスト形式を主体にしながらも、編集支援のためにイメージ情報も取り込もうと試みてきた。この面での解決策の一つが「見出しデータベース」であり「電子縮刷版」である。

基本的に現在では記事データベースはフルテキストデータベースである。これには数々の利点があるが、OCRの誤読率が高い現在の技術水準では、基本的にCTS開始以降の記事しか収録されない。

他方見出しデータベースはCTS以前に遡及が可能である。例えば朝日新聞社は1945年以降の記事のデータベース化をし、95年に「戦後五〇年 朝日新聞見出しデータベース」をCD-ROM 5枚組で発売した。これはCTS以前の記事もデータベースで探したいという調査部の要望に応えたものである。したがって見出しデータベースは縮刷版と組み合わせることで、フルテキストデータベースの欠如した期間を補完し得る。

また日立製作所と読売新聞社、丸善は1993年8月に新聞縮刷版のCD-ROMを開発した。94年1月以降読売新聞の電子縮刷版が年間24枚のCD-ROMで発行されている。

このように電子縮刷版や見出しデータベースは、データの遡及にもイメージ情報の入手にも有効な手段となる。

#### 4. 原紙、新聞切り抜き、縮刷版、新聞記事データベースの相違

2.では新聞切り抜き、縮刷版、新聞記事データベースが、記者の編集支援を契機に生まれたという、それらの共通項に注目した。この4.ではむしろ原紙を含めたそれらの相違に着目したい。

まず縮刷版や新聞記事データベースが、原紙のほぼ正確な再現であるものの、原紙とは異なる点を確認したい。

原紙には訂正やお詫び記事が載るが、多くの新聞の縮刷版には載らなかった。なぜなら訂正済みの記事が掲載されていたからである。しかし朝日新聞社では1989年10月の縮刷版以降訂正せずに、間違った記事そのまま載せる方針に転換した。今では大抵の社の縮刷版は原紙を忠実に再現している。他方原紙は早版と遅版とで大きく違う場合もあるが、縮刷版は原則的に東京本社の最終版に統一される。

コンピュータでデータベース化された記事も原則東京本社の最終版に依拠するが(朝日新聞社「聞蔵」は89年より大阪・西部・名古屋本社のニュース面も収録している)、縮刷版と異なり字句などの誤りは修正され易い。また訂正記事が後で出た記事には、修正は施さなくても訂正記事を張り付けたり、訂正の参照部分を明示するという点で、現行の縮刷版よりも正確性への志向がみられる。だがこれは新聞社による方針の違いがある。またしばしば見出しは書き換えられる。このように修正がなされる一方で、著作権上の理由から新聞社の外部の評論家や学者などの寄稿した一部の文章が欠落する。

このように収録記事の範囲や記事内容そのものや見出しについて、原紙と縮刷版と記事データベースとで、相違がある。そして社による相違もあるが基本的には記事データベースでは事実の正確性が求められ、それに応じて若干の修正があるといえる。

また別の次元の相違もある。

本誌の44巻4号に掲載された、日本経済新聞社、石井昌之の論文は、原紙と記事データベースの相違を次のように記す。

「新聞では記事はその日の紙面の中で関連付けられ、

あるいは見出しの大きさと軽重が判断されて読者に読まれるのだが、同じ記事がいったんデータベースの中に収録されてしまうと、新聞紙面の中で記事の関連性は無視され、扱いの大小もまったく意味を持たなくなる。全ての記事が情報として並列な価値を持つ。その記事の重要性を判断するのは新聞の作り手ではなく、データベースの検索者ということになる」<sup>3)</sup>。

当然ある記事がどの面にあるかで、その記事の意味づけ、その記事に与えた新聞社の価値づけが異なる。また先述のように見出しは書き換え得るが、書き換える前の原紙においても、見出しの大きさはその記事の価値づけに関わる。

そのような価値づけは紙面のなかに記事が組み込まれているから生じるのであって、データベース化された記事は紙面から独立した1つ1つのレコードと化し、価値づけから解放され、「データベースの検索者」の主体性がより発揮される。

しかし価値づけは、単に整理部記者の色眼鏡によってだけではなく、その時代の空気とか、記事を支える全体のコンテキストによっても決まる。つまり価値づけの残る原紙は、新聞社の立場を読者に押しつけるものである反面、読者はそれによって時代背景を明確にさせ得る。

当然この問題は、テキスト形式の通常の記事データベースの問題であって、電子縮刷版などのイメージ形式のデータベースの場合、あるいは見出しデータベースのように紙媒体の縮刷版を常に参照することを前提としたデータベースの場合、媒体上の価値づけは残される。しかし利用者の意識は周りの記事には向かないので、通常テキスト形式のデータベースともさほど変わらない。

他方、原紙及びそれをそのまま素材として切り取った切り抜きの場合、文脈や価値づけを十分に意識させる。例えば人間の脳は左脳がテキスト情報を、右脳がイメージ情報を司るが、右脳の記憶は鮮明かつ情報量も多いとされる。つまり記事を読んだ際の、周囲の状況やそのときの時代の空気、そういった文章化できない諸々の情報が原紙や切り抜きを読むことで思い浮かぶ。

以上を踏まえて理念型的な類型化を試みよう。

原紙には誤りも多いが、歴史史料としての価値がある。記事の置かれた社会的文脈が、見出しや紙面構成から伝わるし、その意味で報道姿勢の検索に適する。他方テキスト形式の記事データベースは、誤りは極力修正や訂正箇所への言及によって正され、正確性にその商品価値がある。また検索する側は報道姿勢よりは事実そのものの情報を求める。この両極のなかで最も

原紙に近いものが新聞切り抜きであり、次に縮刷版が位置づけられよう。他方電子縮刷版はより記事データベースに近い。

## 5. 新聞のデジタル化の展望

### 5.1 日経テレコン21について

そこで次に新聞のデジタル化の現状と展望を見て、記事データベースと切り抜き業務との関係を知る手がかりを得たい。まず「日経テレコン21」を瞥見したい。

「日経テレコン21」は「電子新聞」をコンセプトに1997年1月に開始された。ここでは双方向性を重視し、利用者の要求に添ったサービスを目指している。

「速報ニュース」では1日1,500本のニュースがジャンル別に表示される。これは日経4紙の記事総数800の倍近い数字である。

またこの「速報ニュース」をニュースバリューに応じてレイアウトしたのが「ハイパーニュース」である。なお「ハイパーニュース」は現在では「ピックアップ」に代わり、それは、利用者の検索履歴に応じて、自動的にその利用者にとってのニュースの重要度を判断し、その記事をトップページに表示する機能が付いている。

「今日の新聞」では日経4紙の最新版をいち早く読める。「新聞・雑誌検索」では新聞記事に関しては、日経4紙の他に、全国紙、ブロック紙、専門紙等20数紙の記事が検索可能である。また雑誌記事については日経PB社の雑誌や「エコノミスト（毎日新聞社）」の記事が検索できる。

以上から「日経テレコン21」は蓄積データベースに速報的なデータベースサービスである選択的情報提供(SDI=Selective Dissemination of Information)への志向を加えたサービスであるといえる。ただし当面新聞記事のサービスにおいて蓄積データベースと選択的情報提供とは連続していない。速報ニュースの最新の情報は古い情報に上書きされる。他方蓄積データベースは基本的に新聞記事なら最終版で構成され、選択的情報提供で登場したその折々の最新ニュースは残されない。しかし将来情報の蓄積能力と検索機能の向上する暁には、新聞記事のサービスにおいても選択的情報提供と蓄積データベースとは連続する。

また紙媒体の新聞のもっていた価値づけ機能を「ハイパーニュース」は保っていたものの、現在では「ピックアップ」という形で、新聞社の価値判断を読者に押しつけずに、利用者個々の価値判断をコンピュータが履歴から推測する形に、変化してきているといえる。

### 5.2 全般的な展望

以上見た「日経テレコン21」の例を踏まえて、新聞のデジタル化の意義と問題点を考えよう。

メディアのデジタル化の進展に伴う、メディアの融合・複合化の流れのなかで、新聞社の速報ニュースや選択的情報提供とテレビ局のデータ放送・文字放送との違いが曖昧になる。さらに電子新聞の発展に伴い、日々の新聞報道が速報や選択的情報提供に連続してくる。

また「日経テレコン21」の「速報ニュース」の1日1,500本という数字は、日経4紙の記事総数800の倍近い数字であることから、「速報ニュース」はニュースの選別作業をあまり経ていないものといえよう。速報や選択的情報提供が、このように詳細な情報を「垂れ流す」機能に専心するならば、通信社の業務と近くなる。実際合衆国のニュース検索サービスで最大の利用者を誇るダウ・ジョーンズ社は、新聞発行事業とニュースの通信業務を併せて展開しており、それが検索サービスに成功した理由であるといわれている。

新聞はテレビに較べて遅延性のメディアである。しかしテレビとの区別が曖昧になり、データベースと連続化し、さらにインターネットとも繋がるにつれて、新聞が、以前に比して速報性のメディアとしての側面を増やすと予想される。

さらに新聞記事とデータベースの目的の違いによって、データベースとしての目的に新聞社としての論評的な使命が従属する懸念が生じる。

新聞社系以外の商用データベースでも新聞記事を基に形成されるものがかなりある。新聞記事が客観的かつ正確であるという信頼があって初めてデータベース源として転用される。また新聞社が記事データベースを公開することでその社は事実強いという世評を強め、そのことが逆にまた記者の報道姿勢を客観的なものに向かわせる。

また論評的報道は出来事の成りゆきから身を引き、じっくりと考えることで成立するが、新聞が速報メディアに変わると、そのような報道姿勢も弱まる。

そしてますますジャーナリズムとしての新聞の役割は後ろに引き下がらざるを得ない。

このように新聞のデジタル化と記事のデータベース利用は、報道をより速報的かつ客観的な方向へと促す。

## 6. デジタル時代の新聞切り抜きの意義

このようにデジタル化時代の新聞は、記事が客観的なものになるし、記事の価値づけも利用者個々の手に委ねられる、あるいは個々の検索履歴からコン

コンピュータが判断するものとなる。

このことは受け手の選択の自由が十分に発揮される証しとして肯定的に評し得る。しかし価値づけは時代状況や歴史的な文脈のなかでなされる。価値づけを廃した電子新聞ないしはデータベース上の記事は、個々の記事がもっていた文脈をあえて捨象したものと化す。4. でみたように、新聞原紙上の記事とデータベース上の記事は、基本的に別物と考えるべきである。

では、なぜ記者や研究員たちは編集支援のために縮刷版やデータベース上の記事を、求めに応じて調査部や資料室から提供して貰えるのに、切り抜きを好むのであろうか。

これは見出しなどがそのまま残されている切り抜きにおいて、記事に対する価値づけ、記事の出た文脈が分かる点が大きいのと思われる。しかしそれならば、テキスト形式のデータベースの場合に、この問題は顕在化するものの、電子縮刷版、さらに紙媒体の縮刷版といった、4. の最後に示したタイプのなかで、より原紙に近づくものでは、その問題は顕在化しないという反論もあり得よう。

それはその通りである。しかし縮刷版が老眼鏡の世界になっている熟年記者には読み辛いという問題以上に、縮刷版や電子縮刷版は原紙のもっている肌触りや感触を、伝えきれないという問題もあると思われる。

新聞を隅から隅まで読む人は少ないものの、第1面や社会面あるいは自分の守備範囲の記事は、記者や研究員などの広い意味での書き手や、図書館の切り抜き利用者などの読者公衆ならば、多くが通読しているものと思われる。たとえ少しずつ日に焼けて変色しても、原紙の感触は、その記事を読んだ際のおぼろげなイメージ情報の記憶を蘇らせてくれる。そのような記憶とともに、その記事の時代状況が浮かび、それとともに、見出しなどのその記事の価値づけや記事の背景や文脈も明確になる。人間の右脳記憶は、イメージの記憶であるが、テキストを記憶する左脳よりも、情報量は圧倒的に多いとされる。

この問題は新聞のデジタル化だけの問題に留まらない。近年電子博物館や電子文書館が整備され、歴史的な史料の多くがデジタル化されて、遠方からもアクセスできる。他方、実証史学の世界では、一次史料つまり未公刊の史料を基に論文を書いて初めて独自の研究成果とみなされる。デジタル化されたり、翻刻された史料を用いて、当たりを付けることは是とされるが、そのみに頼って現物を確認しないことは戒められる。

あるいはある歴史小説の大家は、とにかく描くべき事件があると、その現場を旅して、その場所に立って

みることに努めていると語っていた。それによって膨らむイメージを大切にしたいと語っていた。

また書誌学者は、書物をその書かれてある中身の情報（いわばテキストとしての情報）よりも、形態や活字など、書物という媒体そのものの情報の方に、着目して、書物を研究する。

どうしてもデジタル化が進むと、情報の核心部分であるテキスト情報だけが抽出されがちになるし、そうではなくイメージ情報も併せて検出可能となる場合には、今度はそれで媒体そのものの情報性も十分に伝わると勘違いされてしまう。

やはりデジタル化が進んでも、媒体そのものの情報性は伝えられないし、オリジナルの媒体そのものに歴史的な文脈や背景を伝える力があると考えられる。

その意味で新聞切り抜きの使命は、デジタル化時代にこそ大きく存在すると考えられる。なお、本稿では図書館における切り抜きの問題よりも、まず新聞社調査部と記者に関する問題を考察したが、当然、図書館での問題も、以上の記者の問題に準じて考えられる。

2. で示した小野秀雄と小山栄三の見解の相違でいえば、小野の主張は新聞記者や、図書館利用者でも物書き的な存在の人を念頭に置いて、小山の発言で意識されているのはより普通の読者公衆であろう。そして価値や文脈から解放されていく、切り抜きから記事データベースへの歴史的流れは、小野の見方から小山のそれへの変遷でもあった。その延長上に国立国会図書館の切り抜き業務の廃止も位置づけられよう。しかし「知る自由の保障があってこそ表現の自由は成立する」という「図書館の自由に関する宣言」（日本図書館協会、1954年採択、1979年改訂）を見ても分かるように、図書館の利用者の権利は、利用者が表現者となる可能性を前提としながら、表現の自由と表裏一体のものとなっている。つまり記事の書かれた背景や価値をも読み解き、新たに自分が表現者となるような図書館利用者を、「図書館の自由に関する宣言」は求めているであろうし、そうであれば小野のビジョンこそが今必要とされるであろう。その意味で図書館においても新聞切り抜き業務の継続は大切であるし、中止された場合であっても過去の切り抜き資料は貴重である。

#### 参 照 文 献

- 1) 小野秀雄. 新聞の史的価値とその利用法. びぶろす. Vol. 1, No. 3, p. 3-5 (1950)
- 2) 小山栄三. 世論・新聞・切抜. びぶろす. Vol. 1, No. 3, p. 2 (1950)
- 3) 石井昌之. 日本経済新聞記事データベースの品質管理. 情報の科学と技術. Vol. 44, No. 4, p. 176-182 (1994)

**Special feature : Topics around digitization of newspaper. The Clipping Works in the Digital Age,** Yoshihiro GOTO (Institute of Library and Information Science, University of Tsukuba (1-2, Kasuga, Tsukuba-shi, Ibaraki 305-8550))

**Abstract :** In the digital age, people tend to use databases of articles of papers, and they respect the reports of the facts, not the reports of the values. The articles on the databases release from value judgments whereas the articles on the papers are restrained by them. So the articles on the databases contribute the promotion of the interactivity between the sender and the receiver of communications. But the articles on the databases lose the contexts of the articles. The reason why the clipping works exist in the digital age relates to the estimation for the contexts and the values.

**Keywords :** clipping / database of articles of papers / value / context / information of image