静岡県近代新聞記事テキスト化事業計画書

平成17年12月

近代新聞記事検索WEBサイト http://www.npoabc.jp/

入力ボランティア申し込み https://www.npoabc.jp/vregist/entry.php

エイジングブライト倶楽部公式サイト http://www.npoabc.or.jp/

特定非営利活動法人エイジングブライト倶楽部

静岡県内近代地方新聞テキスト化事業計画の概要

1.近代新聞記事テキスト化事業の意義について

(1)静岡県における近代新聞

民友新聞、静岡新報の2紙を中心として、明治初年~昭和16年までの間に発行された地方新聞は、静岡県立中央 図書館に収蔵、公開されているマイクロフィルム化されたものだけでも10万紙面を超えています。これらの紙面からは江戸から明治への急激な社会諸相の変革期や、近代国家としての体裁を急速に整え、やがて暗い戦争の時代に至るまでの間の、郷土庶民の生活、文化が如実に窺える貴重な資料です。

(2)既に行われている新聞資料化の現状

これら新聞資料は、既に静岡県県史編纂事業の中で画像データ化と見出しデータの採取が完了していますが、全ての記事に見出しが振られてはおらず、また見出し語抽出という考え方そのものが、一定の拾い出し基準に基づかざるを得ないという限界を持っており、その意味では、10万もの膨大な紙面から自由自在にあらゆる可能性の記事を探し出すというには未だ不自由な状態におかれているといって良いと考えます。

このような状態では、せっかくの近代史の貴重な資産も、一部の研究者のみの範囲での活用(研究者にとっても不自由な活用ですが)にとどまり、広く誰でもが利用するということには到底なり得ません。せっかくの宝の持ち腐れと言うべきです。

(3)新聞記事の全文テキスト化とWEB公開の提唱と一般的意義

私どもはこれに対し、見出しのみでなく、記事全文のテキスト入力を行なう事業を提起しています。2001年の当NPO設立 準備を進めると同時に、近代新聞テキスト化事業の準備にも着手しました。

記事全文のテキストデータを歴史資産として後代に残すことを第一の目的とし、さらにこれらデータを順次WEBサーバに格納し、WEBサイトでの記事検索を可能とすることを第二の目的としています。

約1年間の準備期間中に、入力仕様、検索システムの開発などを行い、ようや〈2002年5月より、実際の記事入力作業に着手することができました。

6月には、WEBサイトの準備も整い、それまでの間に入力を終えた記事データをこの新聞記事検索サイトで無償公開も開始しました。

何故コストと人手がかかる全文テキスト化が必要なのか? 必須のものであるのか? という疑問に対しては、百聞は一見に如かず、実際に記事検索をしてみていただければと存じます。見ていただければ全文テキスト化することの大きな意義が、実感として感じていただけるものと考えます。

【静岡県近代新聞検索サイト】 http://www.npoabc.jp/

誰でもが、自宅のパソコンで近代期静岡の新聞記事の検索が可能に!

今まで、遠い図書館に足を運び、しかも面倒で目が疲れるマイクロフィルムでしか見ることが出来なかった新聞を、自宅で自由に閲覧することが可能となります。

小学校、中学校の教室で、日本全体の歴史に、地域の歴史を重ねて学習することが可能に! 「余暇事業」の一環として、授業の日本史の大事件にリンクさせて、生徒自身が同時代の地域の歴史や出来事を検索することで、よりリアルな歴史認識を育てることが可能となります。(県教委、新静岡市の教育委員会に提案) 近代史の研究者にとって、研究のための重要なツールとなります!

(4)新聞記事の全文テキスト化事業のNPOABCにとっての意義

当NPOは、元来高齢者や身体に障害を持っている方々など、IT社会の中では弱者となりがちな人々に対し、パソコン技能やIT対応能力を高める研修の場を提供することを目的として設立されました。

しかし、講習のみを行っても、その後に習得した技能を日常的に使用可能な環境がなければ、せっかくの技能もたちまち失われてしまいます。

私どもは、近代新聞テキスト化事業を、パソコン等の基礎技能研修を終えた高齢者、障害者の方々を、□社会の本当の入口にいざなう為のキーとなる事業であると位置付けています。

パソコンの基礎講習を終えた会員は、希望すれば紙面画像CDからの紙コピー複写作業などの準備作業や記事データの入力作業、あるいは入力されたデータの検証作業などに就〈事が可能です。また対価に関しては、無償での作業参加、有給与での参加のいずれでも選択可能としたいと考えています。体力などからみて、長時間の就業は不可能かもしれませんが、まずはちょっとしたお小遣い程度の働きをすることは可能でしょう。

(ただし現時点では原資が不足しているため無償ボランティアのみの参加となっています)

手にする金額の有無や多寡が問題なのではなく、獲得したIT技能、PC操作技能を使って社会に有用な実業を行う事が重要なのです。

就業機会を得て、実務としてIT、PC技能を生かすとともにその作業が、同時に知的な好奇心を充たすことであり、文化的な意義を感ずる事業への参画でもあるという、めったに無い事業こそ、この近代新聞テキスト化事業であるのです。

2.近代新聞テキスト化事業の問題点とNPO参画の意味

(1)近代新聞テキスト化事業の問題点

全ての新聞紙面から、一定の基準で必要な記事を抽出し、その記事の見出しデータのみを収録するという方法より、全文テキスト化が、新聞資産を有効活用する点で、明かに優れていることが明白であるにもかかわらず、この方式を採用する研究機関や団体が全国に一例も無いのは何故でしょうか?

答えは簡単です。費用があまりにもかかりすぎ、公共団体でも企業でも手を出せない方式だということです。

静岡県の近代新聞10万紙面を全文テキストにおきかえるためには、どんなに安く見積もっても、1紙面5000円以上のコストは計算しなくてはならないことになります。全部を処理するのに五億円以上の費用が必要となるということです。

公共団体の予算で、しかも文化事業として単独事業としてこれだけの金額を捻出することはまず不可能です。

仮に民間の企業がこれを行なうとして、ミニマム五億円の投資を、事後に回収できるかどうか?完成した資料としての価値や意義がどれほど高くても、そのことで回収可能とする根拠にはなりません。

(2) N P O 参画の意味と採算性の確保

そもそもNPOの存在意義は、民間でも、公共団体でも手がけることが不可能な、しかし公共にとって有意義である事業を行う主体となることにあります。

ではどのようにして、NPOはその特性を生かし、公共団体でも民間でもできない事業の採算ベースを確保するのか? この方策について以下に触れます。

ボランティアの支援によるコストの大幅な低減

NPOは、コストのかかる事業に際し、広く全国のボランティアに無償で事業に協力することを呼びかけることができるという長所を持っています。新聞テキスト化事業経費の大半は入力者への人件費支払いが占めます。従って、この人件費支払の一定部分を無償ボランティアによってまかなうことができるのなら、この事業の所要経費は大幅に軽減することが可能となるはずです。もちろん、多くの無償ボランティアの方々に結集していただくには、事業そのものが真に公共の福利をもたらすものであり、かつボランティア個々の知的な好奇心、意欲を大いにそそるものでなければなりません。

その意味で、近代新聞記事テキスト化事業は、多数の無償ボランティアを結集することができるだけの、事業としての意義と 魅力を持っているものと確信します。

完成データの有償WEB公開による収益の確保

入力されたデータは、照合修正の作業後に順次当NPOの専用WEBサーバに格納され、格納と同時にWEB公開されます。 WEB検索サイトの利用は原則として有償とし、利用者は当NPOに利用費用を支払っていただくこととします。

有償化は、平成16年年9月に近代新聞検索WEBサイト上で開始しました。

平成17年11月末現在での有償サイトの検索対象記事件数は42,292件にすぎず、完成時の予想総件数(約200万件)からすれば、まだ2%強にすぎず、微々たるものですが、有償登録利用者の皆さんは、一般の有償検索サイトのような「利用料」ということでなく、むしろ、この事業の趣旨にご賛同いただき、入力に携わる代わりに、利用料としてこの事業を財政的に支援していただくこととなり、その意味では一種のボランティアであると位置づけています。

利用料は以下の通りを予定しています。

*民間企業の場合、使用料としてではなく、当NPOの活動全般を支援する賛助会員となっていただければと考えています。 賛助会員は勿論無償でWEB利用が可能となります。

賛助会員 20、000円/年 (1口)・特別賛助会員 50、000円/年 (1口)

新聞WEB有償公開による収益は、前述の1 - (4)の、高齢者や身体障害者の当NPO会員で、有給与での新聞記事データ入力等の作業を希望するものに対する支払い及びWEBサイトの維持費に充当します。

公共団体、企業に対する支援依頼 公共団体への支援要請(助成金)

本事業について、私どもは静岡県をはじめ、県内市町村の文化事業助成金の申請を継続的に行なっています。しかし今のところ、セミナー形式での事業の宣伝として若干の助成をいただいているのみで、事業本体の維持費用の助成金獲得には到っていません。今後も、国も含め広範に助成金の申請を行なってゆく計画です。

民間企業への支援要請

県内の有力な民間企業に対しても、この事業の有益性をご理解いただ〈努力をし、支援をお願いして行きたいと考えています。具体的には、上記 で触れた当NPOの賛助会員に加入していただき、賛助会費を毎年支払っていただ〈形での支援をお願いして行〈予定です。

3. 事業収支の見通し

平成14年から29年までの16年にわたる長期の計画となりますが、収支計画の詳細は別紙資料「近代新聞テキスト化事業 収支計画書」をご覧ください。

民間、公共団体からの支援とWEB検索使用料収入で、平成21年度内には単年度収支を黒字化する決意です。平成20年までの間はNPO会計からの持ち出しとなります。累計では平成26年度内に黒字となる計画です。

いずれにせよ、この収支計画を現実のものとするためには、 公的助成金の確保 賛助会員企業の確保 WEB有償利用者、 、団体の確保が必須であり、NPO会員の総力をあげてこの確保に邁進します。

4. 各新聞のテキスト化作業予定期間

(1)5紙のテキスト化の優先順と完成時期

下の「予定表 - 1」に見るように、まず静岡民友新聞に着手し、次いで静岡新報、最後に重新静岡、大務、函右の順でテキスト化を行い、向こう8年間で5紙全てのテキスト化を完了します。ただし、この予定は予定どおり無償ボランティアが確保でき、かつ前頁3の助成金その他の収入が順調に確保できることを前提としています。いずれかが予定どおりの確保ができなかった場合、8年計画が不足分だけ延長されることになります。

(2)WEB検索サイトの有償化開始時期

有償化は平成16年9月から開始しています。

予定表 - 1 	平成 14年	平成 15年	平成 16年	平成 17年	平成 18年	平成 19年	平成 20年	平成 21年	平成 22年	平成 23年	平成 24年	平成 25年	平成 26年	平成 27年	平成 28年	平成 29年
静岡民友新聞 紙面数 24,054 M24~S16	+						→									
静岡新報 紙面数 65,027 M24~S16								+						—	>	
重新静岡 M9~M17 紙面数 3,945 静岡大務 M17~M25 紙面数 8,867 逐右日報 M12~M18 紙面数 6,409															+	-

5.近代新聞テキスト化作業の実施体制

(1)総合管理

全体管理は、当NPO事務局の近代新聞テキスト化事業責任者が行います。

事業責任者の下に、2名の実務担当者を置きます。

(2)入力仕様の確定

入力仕様書は既に本年4月に第1版がリリースされました。仕様の作成は近代新聞テキスト化事業責任者(以下「事業責任者」)が行いました。今後、入力実務が進行する中で発生する事例に応じた仕様の変更についても、事業責任者がこれを行います。

(3) データエントリーシステムとWEB検索システムの開発

事業責任者の管理の下、本NPOの賛助会員である株式会社パドラック(湖西市、杉本等社長、工学博士)の支援を受けて開発実施しました。

平成14年~16年8月までは簡易版WEBサイトで運営し、16年9月より検索サイトを有償化、また同時期に入力ボランティアの自動登録システム及びボランティア入力サイトのより効率的なシステムへの更新を既に終えています。

(4)ボランティア

年度別ボランティアの確保予定

本事業には多数の無償入力ボランティアの確保が不可欠です。

別紙資料「**近代新聞テキスト化事業収支計画書**」及び4 - (1)のスケジュールに対応する各年度のボランティア必要 員数は表 - 1の通りです。

表 - 1	平成14~17 年の実績	平成18 年	平成19年	平成20年	平成21年	平成22年	平成23年	平成24年	平成25年	平成26年	平成27年	平成28年	平成29年
入力予定紙面数	2,000(1,815) ()はH17.11末	6,000	8,000	8,500	9,200	9,200	9,200	9,200	9,200	9,200	9,200	9,400	9,400
記事件数	44,000	120,000	150,000	170,000	180,000	180,000	180,000	180,000	180,000	180,000	180,000	180,000	180,000
無償ボランティア必要数		109+21	146+29	154+29	167+33	167+33	167+33	167+33	167+33	167+33	167+33	171+34	171+34
—————————————————————————————————————	静岡民友新聞		静阳		その他								

ボランティア参加者の振り分けと各グループの実務

応募いただいたボランティアの方々は、最初は全員入力を行っていただきます。その後一定期間を経て、入力に習熟した方の中から一定数の方に検証作業をお願いすることになります。入力作業担当者数と検証作業担当者数の比率は、常に10:1となるように振り分けを行います。

入力ボランティアの実務

事務局より一定サイクルで郵送される新聞紙面の記事を順次テキスト入力します。入力は、NPO新聞入力専用サイトに準備されている、記事入力と送信を行う専用フォームを使用して、オンライン又はオフライン(プロバイダと定額制で契約している方はオンライン、従量制の方はオフライン)入力を行い、入力が完了した時点で画面の送信ボタンを押してWEBサーバへの送信を行います。送信後サーバ側で、送られてきた記事内容をテキストベースで送信者とNPO事務局担当者に返信します。以上が標準の入力方法ですが、習熟された方の中には、自分専用の入力ツールをアクセスなどのDBで作って、入力された記事群をまとめて送信するなどの方法でされているケースもあります。

ボランティアには、応募後の事務局からの各種説明を受けた後、各個人ごと固有の専用URL、パスワード、ID番号を事務局より通知し、以後これによってWEBサイトに接続して入力作業を行うことができるようになっています。

入力を行う場所

原則としてボランティア参加者はご自宅で、ご自分のパソコンとインターネット通信環境を使って入力をしていただくことになります。ただし、初心者に関しては、下記3ケ所のNPO施設のパソコンと環境を使って入力することも可能です。特に、当NPOの高齢会員や障害の有る会員の方で、当NPO主催のパソコン基礎講習を受講した後に新聞ボランティア登録された方については、上記の3施設において、常駐するNPOスタッフの指導を受けながら入力作業を行っていただくことになります。

初心者研修を伴う入力作業が可能な施設

御門台パソコンサロン 最大7名(7台) 静岡市清水区御門台静岡県障害者マルチメディアセンター 最大5名(5台) 静岡市

*当NPOが静岡県から管理運営を委託されている施設です。

データ検証ボランティア

選抜された検証ボランティアの方々も、原則として自宅にて作業をしていただきます。検証結果によるデータ修正は、データ専用WEBサイトで、該当記事を抽出して実施します。 IDとパスワードは、入力用とは別のものが必要となります。

その他のボランティア

前頁の表 - 1には表記されていませんが、入力、検証ボランティアの増加にしたがって当事業に関する管理業務が、現行の事務局スタッフの手に余る状況となった場合、管理を手伝ってもらうボランティアを募集し確保する予定です。

ボランティアの募集方法

無償ボランティア

無償ボランティアの募集は以下の方法を基本として考えています。

A:一般新聞文化欄での当事業の紹介とボランティア募集状況の掲載

B:ミニコミ誌、情報誌、歴史に関する定期刊行物等への紹介記事掲載

C: 当NPO新聞検索サイトとボランティア募集ページの、各種インターネット検索エンジンへの効果的登録

D:県内高校、大学での単位取得可能ボランティアの一つとして採用されるよう働きかけ(県、市町村教委、大学)

E:県内の公共団体、公民館などに案内チラシを配布し、ロビー等に常置してもらう。

有償ボランティア

有償ボランティアは、以下の基準を充たす希望者の中から、当年度の有償ボランティア給与予算額の範囲内で選抜する。ただし、現時点では財政的に困難のため、平成26年以降に実施することとした。(11頁最終章及び添付の収支計算書を参照してください。)人員は26年度に通年で10名、27~29年度は通年で20名を雇用する。

有償ボランティア採用基準

当NPOの一般会員(55歳以上の中高齢者又は年齢を問わず身体障害者)であること。

当NPO主催のパソコン基礎講習を受講し終わった方であること。

固定した職業についていない方

(5)新聞紙面画像データの確保と管理

本事業でテキストデータ化の対象とする5紙の全ては、静岡県立中央図書館が所蔵する新聞画像CDに収録されている。 本事業の着手に当って当NPOでは、静岡県立中央図書館長より、本事業への図書館所蔵画像CDの借出しと複製の許可及び 複写CDから出力した新聞紙面コピーを入力ボランティアに配布(後日回収)することについての許可を得ています。現在静岡 民友新聞収録CD及び重新静岡新聞CDの複製作業が既に完了しています。

*複製されたCDは、図書館の規定により事業完了後全数図書館に寄贈することとなっています。 作成された記事データに関しては、県立中央図書館はWEB有償検索サイトが開始された時点で、無償で検索する!権利を得る ことになっています。

(6)新聞画像からの印刷出力とボランティア参加者への紙面デリバリー

許可を得て複製された新聞CDから、当NPO事務局にてA3版サイズで紙面の印刷を行い、これをボランティア入力参加者に必要数を随時郵送にて送付します。入力が終わって返送されて〈る紙面についても事務局が保管管理をします。ボランティア参加者は個人によって一定期間中の入力紙面数が大き〈異なるため、ボランティア個人と打合せを行い、郵送サイクルと1回に送付する量を決めています。事務局では、誰にどれだけの紙面が現在渡されているか、あるいは誰がどれだけの紙面の入力を終えたかを常にチェックし、入力の漏れが生じないよう管理しています。

(7)ボランティア参加者への技術的な支援体制

応募当初のボランティアの研修

応募して入力を開始したばかりの初心者ボランティアに対しては、それらの方が県中部地区在住の方であれば、8頁に示したNPOの3つの施設にて、各々の施設の管理者による指導を受けながら研修を兼ねた入力を行うことが可能です。 県東部、西部地区に居住の方々については、これら施設に通うことは不可能ですので、下記 の勉強会を兼ねて、各々の地区で入力に関する研修会を定期的に催す方法で対応したいと考えています。

静岡県の近代史に関する勉強会の定期開催

既に平成14年以降毎年最低1回(平成16年度は国の補助金にて3回実施)勉強会を兼ねたセミナーを実施しています。 場所も静岡市内のほか、浜松市、富士市にて実施しました。

今後も継続的になるべく多い回数実施したいと考えています。講師は、たまたまボランティア応募者の中に、近代史の研究者がいらっしゃったため、この方にお願いしています。(川上努氏 静岡県近代史研究会会員)

入力ボランティア専用のWEB掲示板での相互連絡とQA対応

新聞記事データ入力専用サイトの中に、ボランティアの方々と事務局の間での連絡掲示板を設置しています。入力に際してでて 〈る疑問、処理方法などを自由に書き込み、これに対して事務局担当者が回答します。事務局からの注意事項の連絡についても この掲示板で行います。掲示板の設置は本年7月に入ってからでしたが、今では質問者に対して事務局のみが応答するのでは な〈、ボランティアの方たちどうしで疑問を答え合うというケースも出てきています。

NPO施設での電話によるQA対応

上記 の掲示板では対応できない急ぎのQAなどの場合、NPOの3施設で電話応答を行います。3施設とも平日は必ず事務局 又は管理のボランティアが常駐しています。

6.近代新聞記事テキスト化仕様について

7頁5 - (2)でも触れましたが、入力に関する規則等を定めた仕様書は、4月に第1版がリリースされています。本計画書に添付しましたのでご一読〈ださい。

この事業では、紙面に収載されている記事の全文を、記事内容によって選別することなく、物理的に読めないもの以外は漏れなくテキスト化することを原則としています。一見つまらないと思われる市況の数字が、あるいは公告の商品名や商店の名が、今予想することのできない、将来の研究者の好資料となることもありうることですし、たまたま何かの事件に関連して載せられた個人名が、その者の子孫、関係者にとって貴重な資料と認識されることもあるのです。詳しくは添付の仕様書でご確認ください。

7.新聞WEBによる収益の長期的活用について(平成26年度以降の計画)

別紙資料「**近代新聞テキスト化事業収支計画書**」では平成26年以降ようや〈累計赤字が解消し、年度収益を上げることが期待されます。

平成14年当初計画では有償ボランティアを早い時期から雇用することになっていましたが、実態としては当初期待していた国、公共機関等からの助成金がほとんど獲得できない状態であるため、変更計画(本計画書)では平成26年以降の4年間にようや〈有償ボランティアの雇用が開始されます。

もちろんそれ以前に累積赤字が解消することが明らかになれば、雇用時期を早める事は可能とも考えています。

取らぬ狸の諺の範疇ではありますが、26年度以降に生まれる利益、特にテキスト化が終了する29年度以降について、当NPOの現在の不安定な財政基盤を根本的に支える財源となります。

当NPOが国や静岡県から委託されて実施している高齢者や障害者のための各種事業を、県や国に頼らず、NPO財政の中で行うことは、この新聞事業の収益を原資とすることではじめて可能となるものです。

添付資料

別紙資料 - 1 近代新聞テキスト化事業収支計画書

別紙資料 - 2 新聞入力基本説明書

別紙資料 - 3 新聞入力詳細説明書

参加ボランティアの皆さんと居住されている地域

54名(前ページの平成15年度予定数の50%) 平成15年3月末現在 総計 県内 37名 県外 17名 京都府 1人 北海道 2人 東京都 4人 静岡県東部(富士川以東) 長野県 2人 10人 静岡県中部(富士川以西、大井川以東) 兵庫県 千葉県 2人 神奈川県 2人 18人 静岡県西部(大井川以西) 愛知県 1人 9人 大阪府 2人

無償ボランティア参加者には、有償WEB記事検索サイト利用の無償パスが付与される特典が付与されます。