

## [資料] 機関紙の文字起こしを arsvi.com へ掲載する作業の 記録

### 2022年の『全障連』掲載について

立命館大学先端総合学術研究科

山口 和紀

要旨：

筆者は立命館大学生存学研究所のアルバイトとして、全国障害者運動連絡会議の機関紙『全障連』の文字起こしを生存学研究所 HP(arsvi.com)に掲載をした。本資料はその作業の内容と課題について述べたものである。

arsvi.com へ掲載したのは『全障連』No.01 から No.86 である。これらは別個のページとして掲載された。それらを一覧する目次ページも作成した。

課題として考えられるのは、掲載の方法である。今回は HTML 形式のページに各号ごとにページを分けて掲載したが、この方法のデメリットとして検索昨日の弱さを挙げることができる。

arsvi.com は内部の検索機能を Google 検索の「サイト内検索」機能に依存しており、正確性と網羅性に欠けると考えられる。その解決策として、ワードプレス等の CMS 入稿とし、CMS の検索機能を活用することが考えられる。

キーワード：

全国障害者運動連絡会議、arsvi.com、CMS、文字起こし、アーカイブス

#### 1. 作業の記録

arsvi.com への掲載作業について、その詳細を書く。HP へ掲載されたのは No.01 から、No.86 までである。ただし合併号が合併号の数で数えて 6 号分あり、No.13 というナンバリングで 2 号発行されている。No.1 は 1977 年 3 月 30 日発行、No.86 は 1989 年 6 月 16 日発行である。文字起こしされた延べ文字数は、5,824,338 字であった。

次のような流れで作業した。生存学研究所の仕事が依頼されるメーリングリストに対して、『全障連』のテキストファイルから HTML 化の依頼が流れた。筆者は障害者運動組織 全国青い芝の会の機関誌を arsvi.com に掲載した経験があった★01 ため、この依頼に適任と思い、作業担当者となった。行った作業はテキストファイルを HTML 化することが主であるが、目次ファイル(各ファイルへのリンクページ)の作成も行った。文字起こししたテキストファイルは 1 号

につき1ファイルとして保存されていた。そのため、1ファイルごとにHTML化した。つまり1号ごとにHTMLファイルを作成した。arsvi.comのファイルはテンプレートファイルが用意されているため、それを用いた。まずテキストファイルからプレーンテキストをコピーし、プレーンテキストをHTML化するツール★02で変換した。そののち、テンプレートファイルへ変換されたHTML形式のテキストをペーストした。その他、ファイル番号やアップロード日時、作業日時などを記載した。この作業をすべてのHTMLファイルについて行った。

作業時間としてはおよそ5時間程度となった。作業に時間がかかったのは、ひとつひとつのファイルの作成は単純であるものの量が多い(80ファイルを超えている)ことだった。またツールで変換したHTMLファイルもそのまま使えるものではなく、手作業で改行の修正や見出しのタグ付けなども行った。その後、目次ファイルを作成した★03。目次ファイルへは、1号ごとにHTMLファイルへのリンクを行った。この作業にも1時間ほど要した。関連するメールでのやりとりや、目次の修正、ファイルのチェックなども含めて作業には全体では10時間弱がかかっていることになる。ファイル名は「zsrNNN.htm」とした。NNNには001、002と連続する番号が入る。掲載ディレクトリは../oとした。これは組織についてのファイルを置いているのディレクトリである。

## 2. 今後の課題

筆者の観点から、今後の課題を示す。法的な問題や制度的問題はここでは述べず、あくまでもHPへの掲載と利用における具体的な課題について述べる。本掲載方式の大きな問題はテキストの全文検索機能がないことである。例えば、『全障連』を一次資料として研究しようとする者が、特定のワードの出現回数を調べたいと考えたとする。現在の掲載方式ではこれを調べる方法は1ファイルごとにアクセスして検索することしかない。Google検索でそれと似たことはできるがファイルのインデックスが不正確である可能性があることと、複数ファイルに存在していた場合には検索に出ない可能性もあり、研究で正確性が求められる場面においては適さないだろう。

テキストファイルをそのままアップロードし、ダウンロードできるようにすれば、各々のパソコンの中で検索することができる。ソフトによってはテキストの全文検索も容易である。したがって、テキストファイルをそのままアップロードすることも検討したが、現時点では行っていない。テキストファイルをアップロードせずにテキストの全文検索を行う機能を作ることも可能だが、作業には費用がかかるため行っていない。

したがって、たしかに掲載されてはいるものの、実際の研究利用という側面からは「検索しにくい」という課題があるだろうと考えられる。可能であれば『青い芝』等のすでに掲載されている文字起こしファイルとの縦断的な検索も可能になることが望ましいが、『全障連』の中だけでもテキストを全文検索できるように整備することは意義がある。

解決策としてはワードプレス等のCMS(Content Management System)での入稿にするとい

う方法がある。CMSはウェブ上のコンテンツ管理を容易にするために作成されたソフトのことである。その中でも広く使用されているのはワードプレス(WordPress)というソフトウェアである。レンタルサーバであれば、標準機能として備わっており、容易にインストールもデータベースのセッティングも可能である。

CMSを使った掲載方法であれば入稿や修正が容易になり、なおかつ、検索機能も標準でついている場合が多いため利点がある。この検索機能とは、ワードプレスの例であれば、コンテンツのデータが格納されているデータベースから特定の条件にマッチしたページのみを一覧にすることができる。例えば、『全障連』のタグが付いたページ、かつ、〇〇という単語が入っているページを表示するというようなことが容易にできる。また結果も正確である。ゆえに、ある種のデータベースとその検索機能としての利用ができる可能性がある。これについては今後、実際に利用しながら検討していきたい。

## ■註

★01 廣野俊輔氏が作成した全国青い芝の会機関誌『青い芝』のNo.1(1957年12月)-No.63(1967年12月)までのスキャンしたデータ(PDFファイル)を1号ごとに分割しアップロードした。作業は2021年7月に行った。

★02 用いたのは「テキスト⇄HTML変換」([http://seistreet.sakura.ne.jp/page/doujin/site/doc/tool\\_text2html/index.html](http://seistreet.sakura.ne.jp/page/doujin/site/doc/tool_text2html/index.html))である。プレーンテキストを入れると、HTMLに変換される。

★03 目次ファイルのURL：<http://www.arsvi.com/o/zsr-index.htm>

# Transcription of a bulletin for publication on arsvi.com Publication of the "Zenshoren" in 2022.

Kazunori Yamaguchi

Abstract:

The writer is a part-time employee of the Research Institute for Ars Vivendi, Ritsumeikan University. I published a transcription of the organ newspaper "Zenshoren". This paper discusses the content and issues related to this work.

I have published on arsvi.com, "Zenkakuren" No.01 to No.86. These were published as separate pages. I also created a page of contents that lists them.

The issue I am considering is "the method of publication. This time I published each number on a separate page in HTML format. One of the disadvantages of this method is the weak search function. arsvi.com relies on the "search within site" function of Google Search for its internal search function. This lacks accuracy and comprehensiveness. A possible solution to this problem is to use a CMS submission such as WordPress and utilize the search function of the system.

Keyword:

National Disability Movement Coordination Conference, arsvi.com, CMS, Transcriptions, Archives