

津波デジタルライブラリにおける文献の デジタル化について

今井さやか

Digitalization of documents about Tsunami disasters in the Tsunami Digital Library

Sayaka IMAI

We are developing a Tsunami Digital Library (TDL) from 2003. In TDL, we provided digital books, newspaper contents, videos, CG simulations of Tsunami disaster and etc. by using PC web browser. In 2014, we digitized some documents about Niigata Earthquake in 1964, Genroku Tsunami in 1703, and Movement of villages after Tsunami Disaster in Sanriku.

Keywords : Tsunami Digital Library, Document Digitalization.

1. はじめに

我々は津波に関する文献、新聞記事、津波遡上シミュレーション動画、フィールドワークデータ、津波災害ビデオなどを総合的に管理し、インターネットを通じて広く公開する津波デジタルライブラリ(Tsunami Digital Library: TDL)を開発している[1][2] (図1)。2003年から現在まで続くプロジェクトであり、常にデータの追加が行われ、運営されている。ここでは津波についての情報を得るだけでなく、津波災害に対する防災計画を策定する上でも有効であると考え。2011年3月に発生した東日本大震災後はTDLへのアクセス数が増加し、津波災害への関心の高まりを伺わせた。TDLの大きな特徴は、日本古来より繰り返し来襲する津波災害の歴史資料を、津波研究者の選択を経てデジタル化し、インターネットを通して閲覧可能であることである。日本だけでなく世界各地で発生する津波の現象の記録を閲覧することができ、学術的にも価値あるデジタルアーカイブと言える。本論文では2014年度に実施したデジタル化作業および現在の公開状況について報告する。

2. 津波デジタルライブラリ (TDL)

2.1 文献のデジタル化作業の実施

TDLで公開されている文献は、主に東北大学災害科学国際研究所所蔵のものであり、長年に渡り収集が行われてきた。中には絶版の文献や記念誌など、現在では入手困難な貴重な歴史資料もあり、デジタル化作業においては資料自体を紛失することのないよう注意を払う必要があった。したがって、資料をTDLの委員が直接管理可能であるよう、デジタル化作業の拠点は基本的に大学に置いて実施されてきた。2003～2008年度は群馬大学、2010～2011年度は東北大学にそれぞれ資料のデジタル化の拠点がおかれ、東日本大震災による中断を経て、2014年度は相模女子大学に拠点を移して津波歴史資料のデジタル化作業が再開された。



図1：津波デジタルライブラリ（TDL）トップページ

2.2 文献の選択

TDLの役割は前述したとおり、本、新聞記事、動画、フィールドワークデータなどの津波歴史資料を総合的に管理し、インターネットを通じて広く公開し、様々なユーザが津波に関する情報を得たり、津波災害に対する防災計画の策定の参考資料にしたりすることを可能にすることである。このような資料はほとんどが紙媒体で発行され、一般の目に触れることはなかった。また、TDLで公開されている資料は、歴史的に価値があり、後の津波災害対策策定においても十分参考資料となり得る資料であることを念頭に置いている。そういった津波歴史資料は発行・公開されてから年月がたち、津波研究者などから繰り返し参照されることで歴史的価値を得ることができる。今回の津波歴史資料のデジタル化では、歴史的に価値があるという観点から文献を選択し、2011年3月に発生した東日本大震災に関する資料は除外した。東日本大震災の津波被害に対しては、非常に多くの画像、動画が記録されており、またそれらの多くがインターネットを通してPCだけでなく、個人のスマートフォンやタブレットPCなどで手軽に閲覧することができる。この期間にも国立国会図書館[3]、東北大学[4]などをはじめとして非常に多くのデジタルアーカイブが立ち上げられ、様々な資料を閲覧することが可能となっている。しかし、災害が発生してから4年しか経過しておらず、繰り返し参照される津波歴史資料の選定は難しいと判断した。今回のデジタル化資料を以下に示す。

- 津波災害時の可燃物、火災
 - 消防庁：新潟地震 火災に関する研究、1965年
 - 新潟市：新潟地震誌、1966年
- 元禄津波に関して
 - 千葉県安房郡役所：安房震災誌、1936年
 - 千葉県郷土史家研究連絡協議会：防災災害史一元禄の大地震と津波を中心に一、郷土研叢書IV、1984年
- 集落移動
 - 田中館秀三・山口彌一郎：三陸地方の津浪による聚落移動(1)(2)(3)、齋藤報恩会時報 141～143号 1938年

2.3 デジタル化作業

本研究で行っているデジタル化作業は、資料の1ページを画像としてスキャナ入力するだけでなく、そのページのテキストデータや図表をも別画像として入力する。さらに、テキストデータに対して、章立てに対応したXMLタグを付与することで、キーワードで検索するだけでなく、章や節ごとに、参照することが可能となる。Webブラウザを用いて表示する際には、元のページ構成にこだわらず、章立てごとの表示にし、図表などは文章

中にはサムネイル表示にし、ハイパーリンクで拡大表示することができる。この資料のデジタル化にはいくつかの作業段階を経ている。すなわち、(1)資料を1ページごとにスキャナによりデジタル画像化・図表部分の画像切り出し、(2)OCRを補助的に用いテキストデータの作成、(3)資料の章構成に対応したXMLタグ付文書の作成、(4)TDLデータベース格納、の4段階の作業工程を経て、TDLの公開に至る。図に各工程作業の画面を示す。図2～4は文献「新潟地震 火災に関する研究」からのスナップショットである。図2に示すように、文章中の図や表はその部分をトリミングし、別ファイルとして保存する。

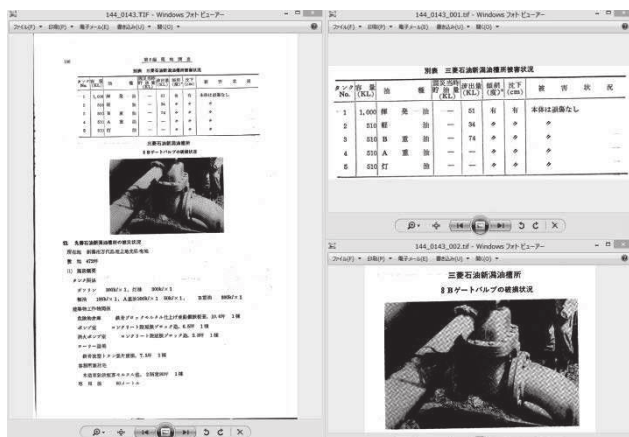


図2：ページごとのスキャナ入力・画像の切り出し

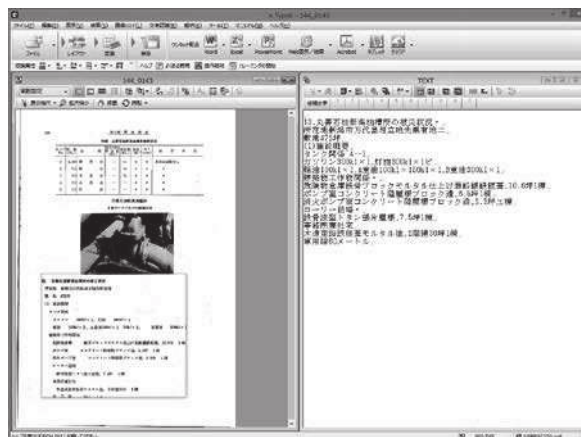


図3：OCR画面

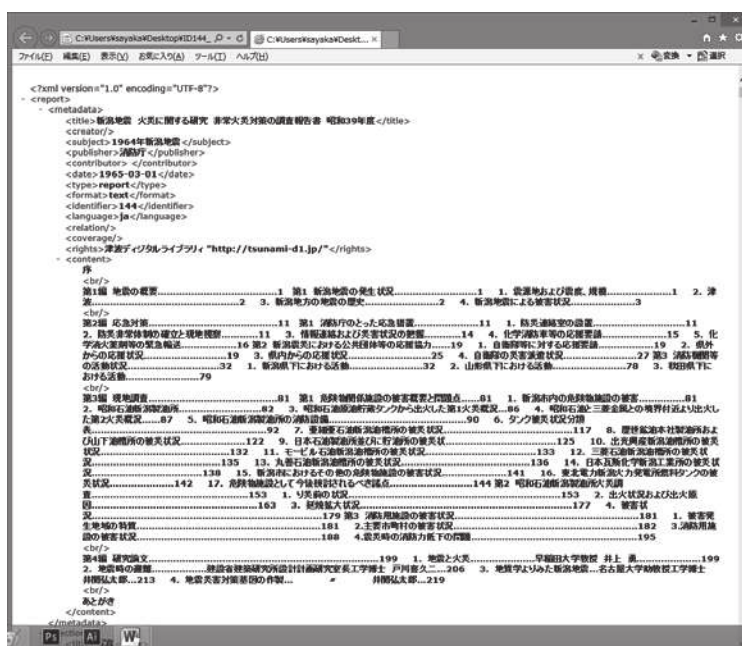


図4：XMLタグ付テキストの表示画面

2.4 TDL データベース

前節の(1)(2)(3)の工程により作成された、XMLテキスト、図表画像データはTDLシステムのデータベースに格納される。データベースを利用することで、津波ごと、文献ごと、キーワードによる検索が可能になる。また、スタイルファイルを適用して表示することにより、章立てなどの文書構成の視覚化が可能になる。デジタル化作業によって、図5にデジタル化された「新潟地震 火災に関する研究」のTDLでの閲覧画面例を示す[5]。



図：TDL での閲覧画面例 サムネイル表示されている図（左図）
本文の表示例（右図）

4. まとめ

2014 年度に相模女子大学を拠点として実施された津波デジタルライブラリの文献デジタル化について報告した。デジタル化したデータは TDL で公開するだけでなく、すべて DVD-R に収録して保管している。また、この DVD-R 集を「津波デジタルライブラリデータアーカイブズ」として国立国会図書館寄贈している。TDL は現在もなお毎月 4000 を超えるページビューがあり、津波歴史資料のアーカイブとして一定の役割を担っていると思われる。次年度以降も同様に継続できるよう努力していきたい。

謝辞

本研究は JSPS 科研費 課題番号 268027 の助成を受けたものである。

参考文献

[1] Sayaka Imai, Yoshinari Kanamori and Nobuo Shuto: Tsunami Digital Library、J. Gonzalo at al. (Eds.) ECDL2006、LNCS 4172、pp.555-558、2006. Springer-Verlag Berlin Heidelberg 2006.

[2] Tsunami Digital Library: <http://tsunami-dl.jp>

[3] 総務省・国立国会図書館東日本大震災アーカイブ ひなぎく： <http://kn.ndl.go.jp>

[4] 東北大学アーカイブプロジェクト みちのく震録伝： <http://shinrokuden.irides.tohoku.ac.jp>

[5] 津波デジタルライブラリからの引用「新潟地震 火災に関する研究」：
<http://tsunami-dl.jp/document/144>