

震災アーカイブの利活用を促進するためのメタデータスキーマについて

東北大学災害科学国際研究所 柴山明寛

東北大学災害科学国際研究所 ボレーセバスチャン

東北大学大学院工学研究科 北村美和子

1. はじめに

2011年3月11日14時46分に国内観測史上最大のM9.0の東北地方太平洋沖地震が発生し、地震動被害と津波被害によって死者19,575人、行方不明者2,577人にもおよぶ多大なる犠牲と甚大な被害をもたらした¹⁾。また、被害は、直接的な被害だけでなく、間接的な被害も日本全域に影響し、さらに海外まで影響を及ぼした。この未曾有の大災害による甚大な被害を二度と繰り返さないためにも、様々な教訓を後世まで残し、今後の防災・減災に繋げて行くことの重要性が言われてきた。東日本大震災が発生した直後から自然発生的に記録を残す震災アーカイブの気運が高まり、様々な機関・団体が震災デジタルアーカイブの構築の動きが見られた。さらに、2011年5月10日東日本大震災復興構想会議において、復興構想7原則の提言が発表された。その原則1には、「大震災の記録を永遠に残し、広く学術関係者により科学的に分析し、その教訓を次世代に伝承し、国内外に発信する」²⁾との提言が発信され、震災デジタルアーカイブの構築の動きがさらに加速することとなった。東日本大震災では、数十団体の震災デジタルアーカイブの構築がなされているが、震災から6年半が経過した現在では、収集された震災記録の利活用について様々な議論がなされている。しかしながら、各団体でも震災記録の利活用が進んでいないのが現状である。

本報告では、震災記録の利活用が進んでいない理由として、震災デジタルアーカイブの根幹となるメタデータ及びメタデータスキーマに着目して、その品質等の問題点について報告する。

2. メタデータとメタデータスキーマとは

メタデータとは、資料のページ数や出典先、提供元、作成日、映像や画像などの解像度など、資料に関係する諸情報を記載しているものである。図書館を例に出すと、書誌情報がメタデータに当たる。メタデータスキーマとは、資料の諸情報を記載するためにどのような項目（ページ数、出典先、提供元など）が必要なのか、その項目の内容をどのように記述するかを決めるものである。記述形式の例としては、資料の提供先の場所情報を記載する際に、場所を緯度経度で記載するか、それとも住所等の地番で記載するかを決めるものである。

メタデータは、資料を検索するために必要不可欠なものであり、メタデータの項目や記載内容が豊富なほど、複雑な検索や資料の絞り込みが可能になる。ただし、同じような資料でかつメタデータに記載されている内容が類似している場合には、資料の絞り込みができなくなり、目的のものを見つけられなくなる。

* About metadata schema for promoting use of earthquake digital archive by Akihiro Shibayama, Boret Sebastien and Miwako Kitamura

3. 東日本大震災の震災アーカイブのメタデータスキーマについて

東日本大震災アーカイブを構築している団体の多くは、国立国会図書館東日本大震災アーカイブメタデータスキーマ³⁾（以下、NDLスキーマ）に準拠して設計を行っている。また、一部ではあるが独自のメタデータスキーマを設計している団体もある。NDLスキーマは、2013年4月に公開され、現在までに2回の修正が行われており、最新版は、2016年10月版となる。また、同時期に英語版が公開されている。2017年現在、メタデータスキーマを公に公開されているのは、NDLスキーマのみであり、その他は、公開がされていない。NDLスキーマに準拠している団体は、自治体が構築している震災アーカイブが多く、研究機関では、東北大学「みちのく震録伝」などもNDLスキーマに準拠している。NDLスキーマの項目数は、262項目があり、書籍や画像、映像など多種多様な資料に対応することができる。しかしながら、最初の設計が2013年と震災アーカイブが構築されはじめた時期であり、震災から6年半が経過した現在では、東日本大震災の震災アーカイブの独自性が強まり、NDLスキーマの項目以外の項目が作成されるようになってきた。例としては、岩手県が構築している「いわて震災津波アーカイブ～希望～」⁴⁾では、「分類（時間）」、「分類（テーマ）」、「分類（カテゴリ）」など、Webサイト上で独自に表現したい項目を設けている場合もある。その他の自治体でも独自に項目を設けている場合が散見する。

4. メタデータの独自項目の問題点について

3章で述べた通り、各自治体では独自の項目を設けている事例がいくつか存在する。この独自項目は、自身のサイトでコンテンツの検索をし易くしたりと有用なものである。しかしながら、国立国会図書館「ひなぎく」⁵⁾等で行っている震災アーカイブのメタデータを収集して横断的検索できるシステムでは、問題が発生する。これは、メタデータ連携する場合のマッピングに関係している。各機関の震災アーカイブのメタデータを収集する際には、収集先のメタデータの項目を受入側のメタデータのどの項目に当てはめるかを定める。国立国会図書館「ひなぎく」を例にすると、網羅性が高いNDLスキーマではあるものの、収集先の項目が独自に作成されており、NDL側に項目自体が無い場合がある。その場合に、いくつかの方法があり、近い項目に当てはめる方法、ある項目に集約する方法、項目自体をマッピングしない方法などがある。これらのどの方法を選択しても、収集先と同様な検索精度を持つことは難しい。今後、震災アーカイブの構築は、他の自然災害でも構築する動きが見られており、今後さらに独自項目が増えて行く可能性がある。

5. メタデータ品質に関する課題について

震災アーカイブのメタデータの項目は、震災関連資料がどのような資料であることを示すために、タイトルやキーワード、作成日時、資料分類などを設けている。その中で、資料のキーワード付けが重要となる。キーワードの付け方は、総務省ガイドライン⁶⁾でも事例が示されているが、どのような語彙をキーワードに用いるかまでは規定がなされていない。そのため、文章以外の写真や映像などのキーワード付けに様々な問題が発生する。例えば、海岸浸食等を防ぐ

ために「消波ブロック」は、「波消しブロック」、「消波根固ブロック」など別名も存在し、また、世間が一般に幅広く知られている名称も存在する。これらの単語は、同じものを指しているが、キーワードを付ける作業者の知識度合いによって付ける単語が異なることがある。一例として、国立国会図書館の「ひなぎく」で検索した結果を表1に示す。「消波ブロック」を用いていることが大半ではあるが、「波消しブロック」や不動テトラの登録商標である「テトラポッド」などで付けられている場合もある。これは、シソーラス辞書によってある程度の語彙の統制は可能である。しかしながら、様々な商品名があるものなどは統一するのが現状では困難である。

表1 国立国会図書館の「ひなぎく」での検索結果

検索キーワード	検索件数
消波ブロック	6,037 件
波消しブロック	269 件
テトラポッド	697 件

※2018年1月に検索した結果

次に、同じ被写体で類似している写真でも機関によって異なるキーワード付けがされる場合がある。例えば、宮城県石巻市の旧門脇小学校は、津波からの避難が的確で多くの生徒や住民が助かった学校である。この学校は、津波によって火災が発生し、その跡も残っている。火災の跡は写真の解像度や角度などで建物の汚れに見えてしまうことがある。そのため、ある機関の写真には、「火災跡」のキーワードが付与されているが、別の機関では「火災跡」のキーワードが付与されないことがある。被写体に対しての知識によってキーワードの付与が左右されてしまうことがある。

その他にも、同一名称同一機関であっても時期によって場所が異なる問題がある。例えば、宮城県の石巻市立病院は、震災当時、沿岸地域に位置しており、津波の被災があった。その後の平成28年に石巻市の中心市街地に移転されている。現在のメタデータの項目では、資料の位置情報が異なるだけで、その他は同一になってしまう。もし、位置情報が付加されていない場合などは、どちらの建物なのか区別がつかなくなる場合がある。その他にも津波で被災にあった小学校や中学校などは、統合や移転などがあり、同様な問題がある。

類似している問題として、同一名称同一形状の問題もある。例えば、宮城県女川町にある「女川のいのちの石碑」などがそれに当たる。2018年1月時点で、16の石碑が別々の場所に建てられ、最終的には21の石碑が建てられる。この石碑は、刻まれている詩の部分が異なるだけであり、一見して判断は難しい。位置情報が含まれていない写真などの場合は、どの石碑か区別がつかず、同じ石碑として扱われてしまう可能性がある。

上記以外にも、ある目標物を撮影しても撮影場所が離れている場合の位置情報の取扱の問題、新聞記事など複数の地名が含まれている場合の取扱の問題、資料の作成日がはっきりしない場合の作成日の決め方の問題、などがある。

6. まとめ

本報告では、震災デジタルアーカイブに関するメタデータスキーマの問題やメタデータの品質の課題をいくつか挙げた。これらの課題は、氷山の一角であり、その他にも様々な問題がある。また、震災デジタルアーカイブのメタデータスキーマは、地震と津波災害の特化したスキーマであり、自然災害全般に対応してはいない。また、利活用を促進するためには、メタデータの内容の充実を図ること及び各資料がどのように結び付いているかを記載することが重要となる。今後、現状の課題を明らかにし、震災アーカイブの利活用の促進に繋げていきたいと考えている。

参考文献

- 1) 総務省消防庁. 平成 23 年 (2011 年) 東北地方太平洋沖地震 (東日本大震災) について (156 報). 2017.9
- 2) 内閣官房. 2011 年 5 月 10 日東日本大震災復興構想会議. <http://www.cas.go.jp/jp/fukkou/> [accessed 2017-12-1]
- 3) 国立国会図書館. 国立国会図書館東日本大震災アーカイブメタデータスキーマ. <http://kn.ndl.go.jp/static/metadata> [accessed 2017-12-1]
- 4) 岩手県. いわて震災津波アーカイブ～希望～. <http://iwate-archive.pref.iwate.jp/> [accessed 2017-12-1]
- 5) 国立国会図書館. 国立国会図書館東日本大震災アーカイブ. <http://kn.ndl.go.jp/> [accessed 2017-12-1]
- 6) 総務省. 震災関連デジタルアーカイブ構築・運用のためのガイドライン. http://www.soumu.go.jp/menu_news/s-news/02ryutsu02_03000115.html [accessed 2017-12-1]