

より詳細かつ相互運用性の高い書誌情報の記述に向けて： 梶尾コレクションを事例として

幾浦裕之^{†1} 永崎研宣^{†2} バイアロック知子^{†3}

概要：古典籍の書誌情報を記述するに際して、資料の多様な性質を、可能な限りうまく反映するためには、高い可塑性のある記述方法が期待される。既存の蔵書検索システムの多くは、しかし一方で、近年よく求められる相互運用性も可能な限り高くする必要がある。この問題の解決策として、TEI (Text Encoding Initiative) ガイドラインが用意されていることはよく知られているが、日本語資料への対応は十分ではない。本発表では、UCLA 図書館所蔵の梶尾コレクションを対象とした TEI ガイドラインの適用方法について検討する。

キーワード：日本古典籍書誌学、TEI (Text Encoding Initiative)、蔵書検索、蔵書史

How to describe the bibliographical Information of Japanese classical texts in detail and conscious of interoperability Description of UCLA Toganoo Collection as an example

HIROYUKI IKUURA^{†1} KIYONORI NAGASAKI^{†2}
TOMOKO BIALOCK^{†3}

Abstract: It is more desirable that we should describe the bibliographical information of Japanese classical texts by a format has plasticity to reflect various aspects of material. Meanwhile, the format inevitably needs the recent burgeoning conscious of interoperability as much as possible. Text Encoding Initiative has possibility to serve for solution to the problem.

However, it is inadequate for TEI to meet the applying to Japanese texts. If we applied TEI to description of bibliographical information of Japanese classical texts, we could extract particular bibliographical information from a large amount of data while we conserve various aspects of each text. The bibliographical study of classical Japanese texts recently needs to develop comprehensively from a broad perspective including pre-modern Japanese Buddhism texts. My research plan to mark up bibliographical data of Toganoo collection of UCLA Library following TEI Guideline. [**]

Keywords: Bibliographical study of Japanese classical texts, TEI, book stock retrieval system, the history of book collection

1. はじめに

現在、既存の蔵書検索の多くは、一定のフォーマットに従って刊行された図書を前提とし、効率的な検索による高度な発見性を提供している。しかし一方で、古典籍においては、そのようなフォーマットでは十分に拾いきれない、より微細な相違や特徴についての情報が個々のアイテムの特徴として重視される場合があり、それは検索においても同様である。たとえば、蔵書印に着目したデータベースが作成されているが、個々の書籍の来歴情報がそれを通じて明らかになる場合があり、書籍流通史をはじめ、様々な観点から有益なものとなっている。

そのような古典籍のアイテムとしての様々な特徴を記載するために、国文学研究資料館では調査収集事業を長年実施してきており、細目調査カード 1 (現行の細目調査カードは 1979 年度以降使用。以下、細目カード) という形で情報の蓄積を行ってきた。この内容は、日本古典資料調

査記録データベース (旧名称、日本古典資料調査データベース) 2 として公開され、(1)所蔵者名、(2)作品名、(3)編著者名、(4)蔵書印等、(5)序・跋・刊記・奥書等の各項目のテキスト検索と細目カードの画像確認ができるようになっていく。

一方、そのように個々のアイテムの特徴を記述することを目指した手法としては、Text Encoding Initiative ガイドライン (TEI-G) が提示する、XML による記述方法がある。TEI-G は、人文学資料としてのデジタルテキストの効果的な構造化を目指すものであり、XML の柔軟性を活かして、記述の自由度を高く保ちつつ高度な機械処理を可能にするマークアップの手法が、分野や研究手法、資料の性質などにあわせて様々な提示されている。書誌情報の記述手法に関してもその章の一つ (10 Manuscript Description 3, 以下、MS 章) を割いており、写本をはじめとするアイテムの特徴の記述が有益な資料に向けた様々なエレメント・属性を

^{†1} 早稲田大学大学院教育学研究科博士後期課程
The Graduate School of Education of Waseda University Ph.D Student

^{†2} 一般財団法人人文情報学研究所 首席研究員
International Institute for Digital Humanities Senior fellow

^{†3} UCLA ライブラリー 日本研究司書
UCLA Library Japanese Studies Librarian

提示している。

2. 研究の背景

2.1 関連する2つの事例

TEI ガイドラインにおける書誌情報記述手法と日本古典籍の書誌情報記述手法の対応付けに関しては、すでに先行して検討した事例がある。4. この先行事例では、上述の MS 章と細目カードにおける項目の対応づけを検討し、エレメントの対応が完全ではないことを指摘した上で、構造面での相違があることから TEI の書誌情報記述における構造を採用することはできないため、TEI の構造は採用せずに、単にスキーマとして参照する形で細目カードの構造化を目指すことが述べられている。すでに大量に蓄積され、入力内容の電子化も進んでいる細目カードの情報を可能な限り自動的に TEI の構造に変換することには十分な意義があり、無理なく変換するために構造までは継承しないという考え方には一定の合理性があるように思われる。

また、東アジアも含め世界各地の古典籍のデジタル画像を公開しているケンブリッジ大学図書館のデジタルライブラリ 5では書誌情報の記述に TEI を用いており、それを Web サイト上で検索できるだけでなく、TEI で記述した書誌情報ファイル自体をダウンロードすることも可能になっている。ここでは、当該図書館の日本研究司書により TEI マークアップされたもの 6 だけでなく、近年は、立命館大学のプロジェクトとしてデジタル画像を公開する際に作成されたもの 7もある。来歴情報も含めた書誌情報が詳細に記述され構造化されて検索対象となっているため、アイテムとしての特徴を対象として検索することが容易であり、また、ダウンロードして機械処理するといった用途にも有効なものとなっている。ただし、参考文献[a]が指摘する課題に対応したものかどうかは未確認である。

2.2 本研究の目指すところ

TEI ガイドラインは XML の可能性を広く追求したものであり、TEI に準拠するすべてのマークアップの意図に完全に対応できる処理系は存在し得ない。むしろ、マークアップする側としては、処理や表示を必ずしも意識せず、あくまでも記述したい情報を可能な限り構造的にマークアップをすることが主眼となる。そして、そのマークアップされた情報は、すぐにすべての構造が適正に機械処理され表示されなくとも、いつか処理系がアップデートされてそれらが意図のとおりに表示される時を待つ、ということになる。さらに、それらのマークアップ情報は、機械処理によって抽出され、場合によっては Linked Open Data として提供されたり、あるいはさらに、可視化され、利用者・読者の理解を助けたり、それまで見えにくかった事実を明らかにしたりする、といった形で活用される可能性を有するこ

ともなる。様々な機関や関係者が別々にマークアップした情報がまとめて機械処理可能となれば、マークアップされた来歴情報や記述に関する様々な特徴など、可視化されることで新たな発見につながる可能性は飛躍的に高まっていくことだろう。このことは、近年のオープンデータ・オープンサイエンスの流れにも合致するところである。

一方、そのようにして幅広く機械処理を可能とするような TEI マークアップファイルが各地で作成されるようになるためには、TEI のマークアップに従うだけでなく、そもそも、日本の古典籍の様々な特徴を、TEI のガイドラインのどこにどう適用していくか、ということについて、ある程度の共通のガイドラインが必要となる。たとえば EpiDoc8 や Best Practices for TEI in Libraries⁹ など、すでに、分野や用途に応じたより詳細なガイドラインが作成されつつあることから、日本の古典籍の書誌情報に関しても、そのようなガイドラインが必要であると思われる。すでに TEI 協会では東アジア/日本語分科会が設立されてそのような活動に着手しているが、それにあたっては、多様な文献の様態を前提とした様々な対応が必要である。

このようなことから、本研究では、日本古典籍の書誌情報に関して、TEI のガイドラインを前提としつつ、どのような情報をどうマークアップすべきであり、どの程度までマークアップ可能であるかということを試行し、その成果を共有して、今後の日本古典籍の書誌情報記述の在り方について検討していきたい。さらに、このマークアップを活用した可視化を通じ、詳細な書誌情報の共通化された記述がもたらし得る可能性についても探ってみたい。

2.3 本研究における対象資料

本研究では、UCLA Library 所蔵の梅尾コレクション (Togano Collection) の古典籍を記述対象とする。梅尾コレクションは日本の高野山大学学長、図書館長などを歴任した近代密教学の碩学梅尾祥雲(1881~1953)の旧蔵書で、古典籍と近代以降刊行の書籍からなり、全体としては 342 の書名、冊数としては 968 冊が数えられる。第二次世界大戦後にアメリカの各大学で地域研究が重視された際、日本語・日本研究のために日本語図書が急激に整備された。そのような状況のなかで、UCLA 東洋学部 (当時) に仏教学の講座が新設されるにあたって仏書の充実が急務となったとき、1962 年から 63 年に第二代東洋学部長足利演正が主導して購入したのが、梅尾コレクションである。祥雲の子息祥瑞(1928-1988)が 1955 年以降、高野山ロサンゼルス別院で開教師を勤めていたことがこの購入の機縁となった。祥瑞はのち 1974 年にはボストン美術館に東洋部仏教芸術主任研究員兼図書館長として就任した。

梅尾コレクションは平河出版社から刊行された『梅尾コレクション顕典籍文書集成』(1981)において写本を中心に 66 点が影印化された。また、それ以外の刊本も含めた約

110 の書名の古典籍の基礎的な書誌情報が 1990 年代に UCLA Library 所蔵の日本古典籍全体のなかで調査され、鈴木淳・三木身保子による『カリフォルニア大学ロサンゼルス校所蔵 日本古典籍目録』（刀水書房 2000）として目録化された。しかし、この目録では祥雲以前の古典籍の伝来を示す、表紙に書かれた署名や、蔵書印など、詳細な書誌情報については省略されている。折しも日本では寺院の経蔵調査が各所で進行し、所蔵する典籍、聖教の研究が深化して、従来看過されてきた仏書の典籍にある蔵書印、僧侶の署名が、伝来を示す重要な書誌情報として調査の際に記述されるようになってきている。また 2017 年には第一回日本宗教文献調査学合同研究集会在開催（慶應義塾大学、9月23, 24日）され、個別に進められていた調査の報告が共有されるようになってきている。梅尾コレクションの古典籍にも、祥雲が入寺し、住職を務めた香川県仁尾市の不動護国寺覚城院をはじめ、周辺寺院の寺院名が署名や蔵書印に散見されることから、改めて詳細な書誌情報を記述し、項目ごとの検索が可能であるような書誌情報が必要となっている。

また国文学研究資料館の取り組みをはじめ、在外資料も含めた日本語の歴史的典籍の調査や、国書、漢籍、仏書という従来の枠を越えた書誌学の総合化も注目を集めている。一方で、現在書物の分野、目的ごとに利用されるデータベースは複数あり、利用者は必要なデータベースを把握し、ひとつひとつ利用している状態である。研究の関心に応じた特定の項目のデータの検索や、横断検索などを実現するためには、将来的には書物のジャンルや特性にそった高い自由度を保ちつつ記述し、かつ相互運用性を確保した書誌データを残せる書誌情報の記述の方法を検討、確立することが求められている。

以上のように、梅尾コレクションは詳細な書誌情報の記述、公開が必要とされる在外古典籍であること、祥雲以前の伝来の過程を示す情報を含み、寺院史研究にも接続する要素をもつこと、日本の歴史的典籍のなかでも、国文学の古典籍ではない仏書のコレクションであること、という理由から、TEI の可能性を追求しマークアップの課題を検討するうえで、ふさわしい資料であると考えられる。

3. 日本古典籍としての梅尾コレクション資料のマークアップにおける課題

3.1 書物の構成とエレメントの配列の順番について

まず、TEI を書誌情報の記述、入力に適用しようとする初心者にとっては、書物の構成（表表紙から始まり裏表紙に終わる）とそれに沿った書誌調査の記述のながれに対して、TEI のエレメントの配列される順番が合っていないように感じる点があるかもしれない。例えば、書写奥書、刊

記など記述対象の古典籍そのものの成立年次に関わる情報を入力する<colophon>のエレメントが、形態的要素を記述する<physDesc>以下のエレメントに先んじて配列されている。この点で、細目調査カードが(32)として1刊記・2奥書・3識語を入力する欄を最後に設けていたのとは異なっている。しかし、書写奥書、刊記は対象とする古典籍を初めて手に取った際、表紙や題名に次いでまず確認したくなる古典籍の重要な要素なのであり、それだけに<msContents><MsItem><title>に次いで<colophon>エレメントが配列されているのも、ある面で合理性があるのである。

ただし、この<colophon>は、写本の本奥書や書物の内容としての成立の経緯（版本であれば序等）などを入力するために利用する<history><origin>以下のエレメントと隔たっているため、対象とする古典籍によっては、どこからを<colophon>に入力し、どこまでを<origin>に入力すべきか判断に迷うかもしれない。ここでは、データとしてはやや冗長な面と思われることもあるかもしれないが、書写奥書や刊記、成立年をそのまま転記する際には<colophon>を用い、書写奥書や刊記、成立年を含む当該古典籍の成立に関する情報を総合的に記入するエレメントとして<history>を利用しておきたい。

3.2 ひとつの TEI のエレメントに対して書誌項目が複数ある場合

TEI で古典籍の書誌を記述する場合に問題となるのは、従来の日本古典籍の書誌調査で記述していた複数の項目があるのに対して、TEI の側にはそれらに対応すると考えられるエレメントがひとつしかない、という場合である。これは[a]の先行事例でも TEI と細目カードの対応の問題のなかで第一に指摘していることである。以下に問題となる例をいくつか示したい。

日本の古典籍のなかには、内題をもたないもの、表紙が成立後に付け替えられているもの、ある作品に対して複数の書名を有するものがある。また、版本であれば、見返し、扉、序、目録、巻、版心など、書物の様々な場所に題名が現れることから、<title>エレメントは当然ひとつでは不十分である。そこで、アトリビュートを付与して、<title type="">の""のなかにその題名がもつ様々な属性を示すことができる。本研究の調査では、<title type="cover">は当該典籍の外題 (cover title)、<title type="caption">は当該典籍の内題、<title type="alt">は別書名 (版心題、小口題など) というふうに区別し、アトリビュートを設けた。

また、日本の古典籍は書物の顔である表紙も題の位置、色、模様、材質など多様な情報を含んでいる。この表紙の記述においては、<bindingDesc><binding><decoNote type="">においてアトリビュートを付与することによって各要素を記述することができる。

しかし、問題はこのように type="" のアトリビュートを

付与することができないエレメントの場合である。例えば、罫界線、字高、匡郭を記述するために利用できる<layoutDesc>のエレメントのなかには<layout>ひとつのほかに適当なエレメントをつけたり、type=""をつけることができない。

3.3 合写、合綴された資料である場合

ひとつの TEI ファイルは基本的に一冊の古典籍に対応するように作成されるべきものと考えられる。このような原則のもとで課題となるのが、一冊のなかに複数の作品が書写されている、いわゆる合写された古典籍や、書写年次など成立の異なる作品が一冊に綴じられた状態になっている合綴された古典籍である。国文学研究資料館の古典籍総合目録データベースのように、個々の合写・合綴された作品を子書誌とし、それらを統合する親書誌のデータを作成するという方法も考えられるが、古典籍のなかには様々な作品から抄出したテキストを書写したものなどもあり、子書誌の作成も難しいような記述対象が存在する。やはり、他の別な TEI ファイルとリンクさせたりするなど、一冊の古典籍に対して複数の TEI ファイルが存在するよりは、ひとつの TEI ファイルに一冊の古典籍を対応させたほうがよいであろう。

4. おわりに

かつてネットで古典籍の画像を手軽に閲覧することが困難であった時代には、ネットで公開された書誌情報に表紙、奥書、刊記など最低限の画像が数枚だけでも附属すれば、どれだけ検出された書誌データのなかから必要なものを判断する際の助けになるか、という議論、提言がなされたことがあった。古典籍の全ページの画像公開、利用が一般的に可能になった現在では、これからは古典籍の書誌情報は従来とは違ったものになるだろう（文字情報としての書誌情報はより簡略なものになるだろう）という見方も一方ではある。つまり、従来のテキストとしての書誌情報と、画像としての書誌情報の主従関係が、かつてとは逆転したような状態なのである。

しかし、画像で伝えられる情報にも当然限界はあり、画像から正確な書誌情報を読み取り、記述するには訓練と技術を要するのである。古典籍の画像が公開される際に、そこに正確な書誌情報のテキストが附属することが重要であることは依然として変わらない。そのような最低限の書誌情報のテキストなくして公開された古典籍の画像は、ネット上に投げ出されたような状態なのであり、研究に利用する上では注意を要する資料であるだろう。

大量の古典籍の画像が公開されるなかでますます重要となるテキストとしての書誌情報を記述する際に、その対象に応じた自由度と相互運用性を確保するための手段として、TEI を利用する利点は十分あると考えられる。本発表では TEI のもつエレメントを十分に活用したつもりではあるが、まだまだ日本古典籍の書誌の記述に利用できるにもかかわらず用いることができていないエレメントもあるかもしれない。今後も活用できるエレメントの拡充を検討する必要があると考えている。

謝辞

本研究は JSPS 科研費 15H05725 の助成を受けたものです。UCLA Library での調査に関しては、The Tadashi Yanai Initiative for Globalizing Japanese Humanities の一環である UCLA-waseda リサーチフェローシッププログラムでの研究の成果の一部です。

参考文献

- [1] http://www.nijl.ac.jp/pages/images/card_s.pdf (アクセス日: 2018/04/15)
- [2] http://base1.nijl.ac.jp/infolib/meta_pub/G0000401SCH (アクセス日: 2018/04/15)
- [3] Text Encoding Initiative Guidelines P5, 10, <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/MS.html> (アクセス日: 2018/04/15)
- [4] Kazushi Ohya, Unit-based Scheme Connection Between TEI and Original Scheme To Promote Data Sharing Beyond Cultural Diversities, TEI conference 2014, <https://cpb-us-e1.wpmucdn.com/sites.northwestern.edu/dist/8/162/files/2014/10/Ohya-t5w9ev.pdf> アクセス日: 2018/04/15)
- [5] Cambridge Digital Library <https://cudl.lib.cam.ac.uk/> (アクセス日: 2018/04/15)
- [6] Noboru Koyama, Ise monogatari, <https://services.cudl.lib.cam.ac.uk/v1/metadata/tei/PR-FJ-00734-2014-08-15>. (アクセス日: 2018/04/15)
- [7] Zengxian Li, Hyōsen Kōshi Kego, <https://services.cudl.lib.cam.ac.uk/v1/metadata/tei/PR-FB-00769-00001-2016-08-31>. (アクセス日: 2018/04/15)
- [8] <http://www.stoa.org/epidoc/gl/latest/> (アクセス日: 2018/04/15)
- [9] <http://www.tei-c.org/SIG/Libraries/teiinlibraries/> (アクセス日: 2018/04/15)