NEWSLETTER

令和大蔵経と次世代人文学



https://creativecommons.org/licenses/by-sa/4.0/deed.ja

https://reiwadzk.dhii.asia/

1 「大蔵経」と「正法護持」

菊谷竜太［高野山大学］

2026（令和8）年9月30日。インド・ラダック地方の古刹・マンギュ（Mangyu）僧院を訪れたわたしたちは、壮麗な壁画や塑像とともに書架に大切に収められたチベット大蔵経を目にした。言い伝えによれば、それらは昔かの地より遠く離れた中央チベット・ラサの地で印刷され運ばれてきたもので、途中河を渡る際に事故で一部が濡れてしまったが、その痕は今も残っているという。「ラサ版」と呼ばれるこの大蔵経は、ダライラマ13世・トゥプテンギャンツォ（1876–1933）が発願し1934年に開版されたもので、私財を投じ、自らも校閲にあたった。版木の材料には桜の木をわざわざヒマラヤより取り寄せたと伝えられる。その開版に関して多田等観（1890–1967）は一つの興味深いエピソードを伝えている。

ラサ版の出版が計画されたとき、等観は13世に従来の伝統的経帙版**（図1）**ではなく、より利便性をそなえた活字をもちいた小型の装丁を進言した。しかし13世はこの提案を斥ける。「濁世においてはかえって苦労を重ねて出版したい。正法護持の精神が何よりも肝要である。みだりに簡略を賞賛する考えは正法に対する冒涜である。チベット仏教においてはその方針は採用しない」（多田等観『チベット』岩波書店、1942年、pp. 103–104。一部表記・内容を改めた）。13世が語る「正法護持の精神」が単純に字義通り「労苦を厭わない」という意味ではないことは推して知るべきであろう。想像を逞しくして言うならば、伝統的経帙版が形態的にインドでもちいられた梵夾に由来すること、翻訳・校閲を経て、刻字、紙の準備、印刷、装丁へと至る一連の開版作業そのものがチベットで培われた数多の伝統文化の継承のうえに成り立っており、安易な変更はそうした文化的継承基盤の破壊、まさしく「正法の消失」をも含意したものと言えるかもしれない。

それでは、急速に発展する人文情報学（DH）の恩寵を受ける21世紀の我々にとって「正法護持の精神」とはいかなる意味をもつのだろうか。複数の大蔵経に記されたテキストをTEIによって紐付けたり、IIIFに準拠して経典の写本と版本とのイメージを有機的に繋げてみたり、さらに生成AIをもちいた文献の内容解析も実行できるようになった現在、一見すれば、膨大な情報を瞬時に扱いうるという側面においてはある種の変容と言えるかもしれない。しかしながら、それを単純に「簡略」とは断定しがたい。大蔵経を電子化するにあたっては、個別の大蔵経の編纂過程あるいはパッケージに秘められた仕組みを理解し、さらには利用者の利便性をも勘案しながら情報として落とし込む、複雑で膨大な作業が要請されるからである。SAT（大正新脩大藏經テキストデータベース）に代表されるかような取り組みに続き、東北大学はToUDA（東北大学総合知デジタルアーカイブ https://touda.tohoku.ac.jp/portal/ ）においてチベット資料の電子化を進めている。かつて「正法」を我が国に伝えるという13世と等観との想いのもと将来されたそれらを真の意味で「護持」しうるかどうかは、これからのわたしたちの取り組みにかかっている。



【図1】伝統的経帙版（デルゲ版チベット大蔵経 経典データベースより）
https://touda.tohoku.ac.jp/collection/database/tibet/collection

② デジタル世界におけるテキスト表現

ウィッテルン・クリスティアン［京都大学］

■ はじめに

文字によって継承されてきた文化遺産をデジタル世界へ移行させることは、今の時代の重要な課題である。数千年にわたり、テキストは獣皮、粘土板、竹簡、絹、石、紙など、さまざまな媒体に記されてきた。これらの媒体に共通するのは、文字が物質的な形でそこに存在し、中間的な変換を経ずに直接読み取れるという性質である。ごく最近まで、これはあまりにも自明であり、あえて指摘する必要もなかった。

ところが、デジタル媒体ではこの基本条件が変わる。まず、資料に記されたテキストを何らかの方法で入力し、デジタル情報に変換する。次に、その情報を保存し、必要に応じて画面などに可視的な形で表示する。つまり、入力と表示という２つの根本的な変換が必要になるのである。

この２つの変換は避けられない。したがって、原資料にできる限り忠実な表現を目指す必要があるが、変換された情報と原資料とが完全に同一になることはない。地図が、それの表す土地そのものにはなり得ないのと同様である。

デジタル化には、資料の外観を画像として記録する「Digital Facsimile（デジタル・ファクシミリ）」と、文字を符号化して記録する方式がある。前者は写真による複製の延長に位置づけられる。後者は活字印刷との関連を持ちながらも、デジタル時代に成立した新しい表現方式である。画像は資料の姿を伝え、符号化テキストは検索や分析などの処理を可能にする。

筆者は学生時代以来、文献学的研究と文化伝承の理解を深めるため、テキストの可能性を引き出す符号化の方法を探究してきた。本稿ではこの問題を中心に論じる。ただし、文化遺産の記録を保存するうえで、Digital Facsimile が果たす役割も不可欠である。

■ テキストをいかに符号化するか

テキストの符号化は、単に文字を入力する作業ではない。文字の大きさ、紙面上の配置、文字間の余白など、資料に表れた特徴をどのように捉え、記録するかという判断が伴う。その意味で、符号化には解釈が含まれており、作成されるデータの有用性は、その方法に大きく左右される。

符号化は、おおむね「文字符号化」と「マークアップ」に分けられる。前者は、どの文字を記録するかに関わる。後者は、記録された文字列の構造や特徴を記述する。たとえば、ある箇所が見出しや偈頌（げじゅ）であることを示すのは、マークアップの役割である。

文字符号化は、この数十年間に大きく進展した。歴史資料に必要な文字も、多くの場合、一般的な機器で扱えるようになっている。一方、マークアップには、あらゆる用途を満たす単一の解決策はない。同じ資料でも、伝承の研究、言語分析、一般向けの閲覧では、必要とされる情報や表示の仕方が異なるからである。

そこで重要となるのは、用途に応じた表現へ変換できることを前提に、課題に適した方式を選ぶことである。以下では、筆者が取り組んできた TEI、Mandoku、BKK の３方式を紹介する。いずれもデジタル環境での利用を目的とするテキストベースの形式であり、相互変換も可能だが、その際に情報が失われる場合もある。

■ TEI によるマークアップ

Text Encoding Initiative（TEI）は、あらゆる時代と言語のテキストを記述するためのガイドラインを、約 40 年にわたり整備してきた。世界各地の利用者が蓄積してきた知見に基づき、多様な特徴を機械可読な形で詳細に記述できる。そのため、学術研究における標準として広く推奨されている。

TEI では、本文の構造だけでなく、資料に見られる異同や校訂なども明示できる。何をどのように解釈して記録したかを、データの中に詳しく残せることが、その大きな利点である。

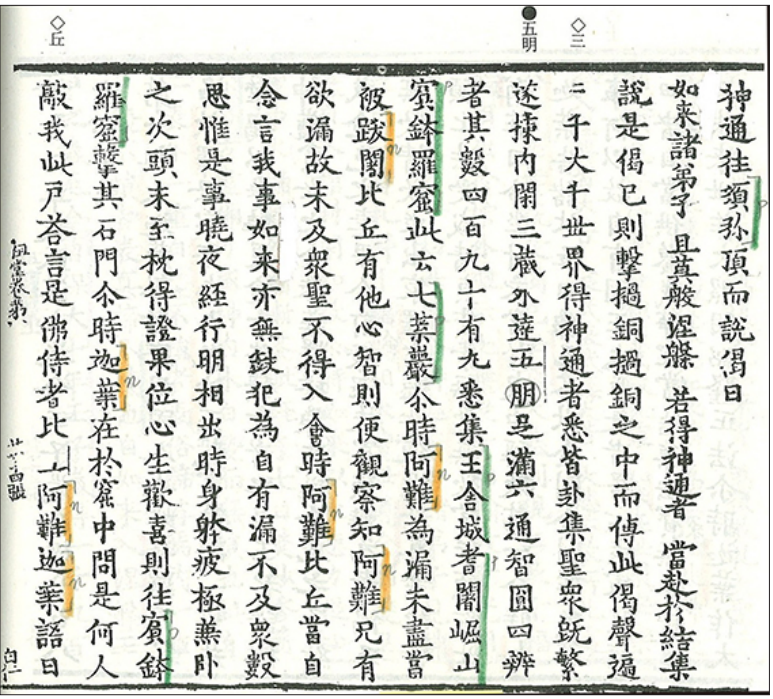
ただし、十分に活用するには、理論と実装の両面で専門知識が必要となる。したがって、組織内に知識を蓄積でき、高品質な研究用データを長期的に作成するプロジェクトに適している。筆者が携わった CBETA（財団法人 佛教電子佛典基金会 https://cbeta.org/）でも TEI を採用した。当初の習得には相当の労力を要したが、長期的には有益な投資であった。

■ 最小限のマークアップ― Mandoku

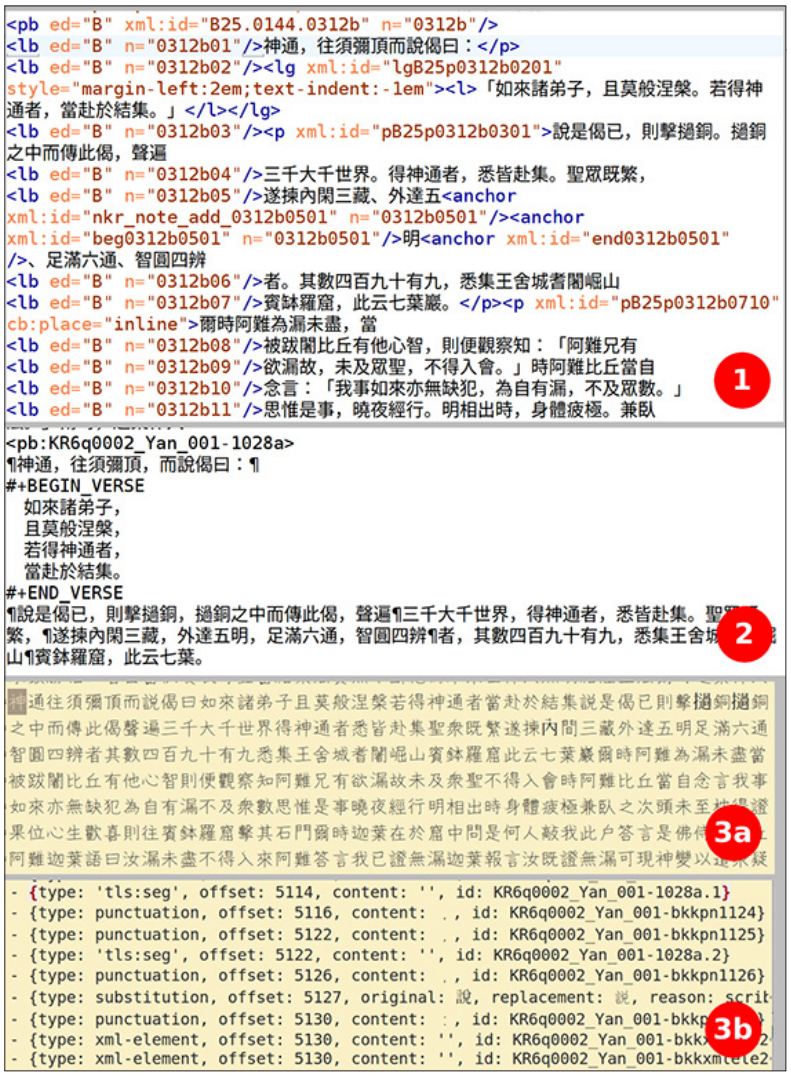
TEI の詳細な記述能力は、その複雑さとも結びついている。この複雑さは、文献研究者がデジタル・テキストを扱う際の障壁にもなってきた。そこで「漢籍リポジトリ」(https://www.kanripo.org/) では、Mandoku（https://github.com/mandoku/mandoku）という、マークアップを最小限にとどめた形式を採用した。

この方式では、少数の特徴のみを明示し、それ以外は暗黙の情報として扱うか、記述対象から外す。すべてを詳しく記録することよりも、本文を扱いやすい形で提供することを重視するのである。

『祖堂集』**（図②、② Mandoku）**の例では、偈頌であることは明示されるが、「五明」に関する校訂**（図①、５行**



【図 1】『祖堂集』（KR6q0002 ／ B0144）に収められた、摩訶迦葉伝の 1 頁。地名と人名は傍線。



【図 2】『祖堂集』の同一箇所を 3 種類の形式で表現した例
① TEI、② Mandoku、③ BKK（a= 文字層、b= マークアップ層）

目上部）は標示されない**（図②、① TEI）**。ただし、「漢籍リポジトリ」は伝承された形の本文も保持しているため、両者を比較することで校訂情報を復元できる。

Mandoku では、本文を符号化し、Digital Facsimile と対応づけ、注釈を付与することができる。記述できる範囲は TEI より限られるが、簡潔な形式でテキストを利用したい研究者の需要に応えている。両者は、異なる目的と利用者を想定した方式といえよう。

■ 分散レイヤーとしてのテキスト― BKK

TEI と Mandoku は、人間にとっての可読性と機械可読性のバランスこそ異なるものの、テキストを一つの論理的単位とし、エディタで開けるファイルに記述するという前提を共有している。

しかし、大規模言語モデル（LLM）の登場により、こうした前提を見直す意義が生まれた。従来も、テキストを訓練や推論の処理に適した形へ変換することは可能であった。ただし、その処理で得られた情報を元のデータに反映する作業は煩雑になりやすい。

この課題に対応する BKK（後述する Bunkankun の略）形式は、文字の層とマークアップの層を明確に分離する。本文を一本の長い文字列として保存し、従来は本文中に埋め込まれていた諸特徴を、独立した「マーカー」のレイヤーに記述する。各マーカーは、文字数に基づく開始位置と、必要に応じた終了位置によって本文に対応づけられる。

たとえば、ある範囲が偈頌であるという情報は、その範囲の位置とともに別の層に保存される。句読点、改行、空白も本文から分離されるため、本文自体は何千文字にもわたる連続した文字列となる**（図②、③ BKK 形式）**。

このデータを人間が直接読むことは難しい。しかし、保存形式と閲覧時の表現を分けることで、必要なレイヤーを選び、目的に応じた形で表示できる。エージェント型 AI の利用は、プログラミング経験の少ない利用者にも、こうした閲覧アプリケーションを作成する、または各種分析などの可能性を開く。試作版は文看訓 Bunkankun（https://bunkankun.org/）で公開されている。

読者は、本文の伝承史、内容に関する情報、文法的・修辞的構造、他のテキストとの関係など、自分が必要とする部分や特徴を選ぶことになる。同じ本文でも、選択するレイヤーとその組み合わせによって、提示される姿は変化する。

この選択は、単なる付加機能ではない。読書そのものが、その時々の目的に応じて素材を配置し、表現を構成する能動的な過程となるのである。

■ おわりに

デジタル・テキストの時代が始まって半世紀が経過した。われわれは今、デジタル媒体を物質的世界の延長として理解してきた思考の枠組みから、徐々に自由になりつつある。テキストを静的に与えられたものではなく、読者が必要に応じて呼び出し、具体的な形を与える動的な存在として捉えるとき、その可能性は新たな広がりを見せる。符号化の探究は、資料の保存や再現にとどまらず、文化遺産を読む方法そのものを問い直す営みなのである。

3 公開学術集会から――DH の最前線

朴賢珍・王一凡・高須賀萌

本研究課題では、令和大蔵経編纂に関わる様々な要素について国内外で広く議論を行い、仏教学のみならず人文学全体にも資することを目的として、仏教学及びデジタル・ヒューマニティーズに関する公開学術集会を開催してきている。また議論の場を広く公開するために、基本的にはオンライン同時配信も実施している。これらについて、以下に簡潔に報告する。

2026 年 8 月 10 日（月曜）～ 14 日（金曜）［発表］

● TEI 2026: Creating Connections, Unsettling Practices

https://tei2026.tei-c.org/

TEI コンソーシアムの年次大会である TEI 会議が今年はカナダ・バンクーバーで開催された。この会議は特に TEI ガイドラインやその応用に関する話題が取り扱われ、XML 技術を介した DH の実践的な報告が集まることを特徴とする。

会場となるブリティッシュ・コロンビア大学のバンクーバー・キャンパスが先住民族の領土に立地することを背景に、開会挨拶と 2 日目の基調講演はいずれも先住民文化をテーマとした内容となった。講演で Jon Corbett 氏は、先住民の概念体系に立脚したデジタルアート制作や計算機インターフェイス開発に関する自身の独創的な取り組みについて紹介した。

セッションでは今回、日本からのポスター発表が目立ち、合計 8 件に上った。本研究課題関連では持田玲が『正治後度百首』の TEI マークアップ、および王一凡が SAT テキストの悉曇符号化と、規格化が見込まれる言語変種コードについて計 3 件を出展した。口頭発表は日本からの発表 4 件中 3 件が本研究課題関連であり、高須賀萌が黄表紙の IIIF 記述モデル開発について、永崎研宣が東アジアテキストの電子化推進の取り組みについて、王が琉球危機言語辞書の TEI Lex-0 マークアップについてそれぞれ紹介した。

研究発表と並行して、TEI の各種ワークショップ、SIG（Special Interest Groups）会合や次期ガイドラインとなる TEI P6 に関するパネルが催されるとともに、北米の長期プロジェクト Women Writers Project をテーマとする討論が開かれ、Susan Brown 教授による最終日の基調講演では、Linked Data によって主流社会の知識体系からマイノリティの知の解放をめざす長年の取り組みが語られた。

このように、共通の技術基盤を軸として多様な視点を包括するコミュニティの存在が実感される百花繚乱の内容であり、来年の東京開催に向けても期待が高まる。

2026 年 9 月 1 日（火曜）～ 4 日（金曜）［発表］

●第 36 回 EAJRS（European Association of Japanese Resource Specialists）会議

https://www.eajrs.net/conferences/2026-leiden

ヨーロッパを中心に日本に関する資料や情報を扱う専門家達が集まる国際団体である EAJRS の第 36 回年次大会が開催された。

本年はシーボルトの江戸参府（1826 年）から 200 年目の節目にあたることから、シーボルトの収集資料を数多く所蔵するライデン大学図書館がホスト機関となり、同大学内の会議施設を会場に開催された。

基調講演では、本会議のテーマにもある「旅」に関連した発表がなされた。渡部泰明館長（国文学研究資料館）は、富士山の煙がどのように詠まれてきたかを文学史的な背景を踏まえて考察した。また、Joanne Bernardi 名誉教授（ロチェスター大学）は、絵葉書をはじめとする 20 世紀の視覚・物質文化資料の収集成果と、それらを「目的地としての日本」というテーマのもとデジタルアーカイブ化し、研究や教育に活用する取り組みを報告した。

さらに、シーボルト関連資料についても、ライデン大学図書館、フランス国立図書館、東京大学図書館など日欧双方の機関から報告があり、活発な情報交換が行われた。

デジタルアーカイブの動向としては、国文学研究資料館、東京大学史料編纂所、東北大学といった日本国内の機関による資料公開状況の紹介に加え、検索ツールの統合化に向けた傾向が窺えた。あわせて、複数の機関から欧州におけるローマ字表記の課題が提示され、実務的な議論が交わされた。

本研究課題関連では、永崎研宣が東アジアテキストデータの構造化の取り組みについて報告し、高須賀萌が黄表紙を具体例とした画像とテキストを組み合わせた複合的な構造化手法を示した。

最後のパネルディスカッションでは、急速に普及する AI との向き合い方について、研究・教育での活用と倫理的観点から熱い議論が展開され、今後の国際的な研究交流と情報共有の発展につながる、きわめて有意義な大会となった。

2026 年 9 月 11 日（金曜）～ 13 日（日曜）［共催］

● JADH2026 “Whose World, Whose Data? Sustainability in Digital Umwelts”

https://www.jadh2026.kyushu-u.ac.jp/

九州大学伊都キャンパスで開催された JADH2026 では、「誰の世界、誰のデータか」を問い、デジタル環境におけるデータの持続可能性をテーマに、データの長期保存や技術的な維持にとどまらず、データが生成・利用・再解釈される社会的・文化的文脈や、その管理と継承に伴う倫理的課題についても幅広い議論が交わされた。

本会議に先立つ 11 日には、プレイベントが開催された。そのうち、本特別推進研究が共催に加わった“Reading, Navigating, and Experiencing Documents: From Bibliography to TEI, IIIF, VR, and the 3D Scholarly Edition”では、ヨーロッパの初期刊本や中世写本、東アジアの写本・版本、とりわけ仏典を対象に、多層的・複合的な構造をもつ文書をデジタル環境においていかに記述し、再現・体験しうるかが検討された。このなかで研究分担者の永崎研宣は、巻子本から TEI/IIIF に至るまで、東アジアの仏典がどのように読まれ、たどられ、利用されてきたかという観点から報告を行った。

12 日からの本会議では、基調講演として、カナダのブリティッシュコロンビア大学の Mark Turin 准教授が、ウイグルの文化的知識の保存と継承を目指す“Rahile Dawut Uyghur Culture Archive”の構築に向けた取り組みを紹介し、メタデータの設計やコミュニティ主体の運営・管理の重要性を論じた。また、一般発表やポスターセッションでは、AI を活用した歴史資料の文字認識や構造化、文学作品の計量分析、TEI によるテキストデータの整備、デジタルアーカイブの構築と再利用など、多様な研究成果が報告された。

こうした議論は、デジタルデータを単に蓄積するだけでなく、その学術的信頼性や再利用可能性を確保し、異なる研究環境や次世代の研究者へと継承していくことの重要性を示すものであった。これは、国際標準に基づく仏典テキストの構造化や、校訂過程の記録・共有を進める「令和大蔵経」編纂にとっても、重要な示唆を与えるものとなった。

DH の最新状況をキャッチアップするためのイベント情報は「人文学情報月報」をチェックしてください

奥付

ニューズレター　令和大蔵経と次世代人文学　第 7 号　**発行日**――2026（令和 8）年 10 月 31 日　**編集発行人**――研究代表者　下田正弘
発行――「デジタル研究基盤としての令和大蔵経の編纂」事務局　**URL**――https://reiwadzk.dhii.asia/　**刊行サイクル**――季刊
研究課題――「デジタル研究基盤としての令和大蔵経の編纂―次世代人文学の研究基盤構築モデルの提示」　**研究課題番号**――25H00001

CC BY SA
科研費 KAKENHI