

はじめに

下田正弘・永崎研宣

第2部は、本科研に参画した研究者による論文の集成である。その導入として、各 chapter の内容について、SAT プロジェクトが進む方向性におけるそれぞれの位置づけを試みながら、その概要を紹介したい。大きく「Ⅰ デジタルコンテンツを作る・使う」「Ⅱ 研究基盤を作る」の二つに分け配列した。

まず「Ⅰ デジタルコンテンツを作る・使う」である。

chapter.02 小野基「仏教論理学研究の現在と人文情報学」では、氏が1980年代に開始したインド仏教論理学文献の電子テキスト作成とそれを取り巻く周辺状況、電子テキストから索引を作成する意義、電子テキストに対する校正作業の困難、というデジタル化に関わる技術的問題を前提としつつ、新出サンスクリット写本やその断片の研究状況、現存しないサンスクリットテキスト再構築という、最新の研究状況にもとづいたインド仏教論理学研究の本質的な問題に切り込み、人文情報学的な課題を提起している。ここで示されるリンクの方向性や付記に記された研究活動記録の集積と共有は、まさに SAT-DB が課題とするところであり、ことに SAT-DB とサンスクリット文献とのリンクへの提言は、SAT-DB の今後の方向を示したものである。

chapter.03 船山徹「文字検索のさらなる地平に向けて—文字列の散在的一致を網羅するために」は、テキスト検索の新たな手法への期待を、丁寧な事例紹介とともに提示する。ここで求められる検索は、Word2vec 等の手法によってある程度実現できる可能性はあるものの、たしかに単純に適用しただけでは、類似の候補が大量にヒットしてしまい、その中からさらに人が目で探さなければならないという状況に陥りがちである。この問題をどう解決するかは、今後の一つの大きな課題である。氏の課題とする「一致」は、現在のほとんどの検索機能が対応するところの表層のレベルの一致を超え、一段深層にみられる一致の発見であるため、情報工学にとって、一段と精緻な分析を要求するものと

なるだろう。その課題の解決は、いわゆる「暗黙知」としてこれまで処理されてきたプロセスをかつてないかたちで顕在化させるものとなり、人文学の発展にとって、極めて重要な一歩となるだろう。

chapter.04 八尾史「^{こんほんせついつさい う ぶりつ}仏典の切れはしを読む方法—「根本説一切有部律薬事」新出サンスクリット写本の研究とデジタルデータ」では、断片化された写本群の中から新たに発見された典籍をつなぎあわせて解読することについて、その手順の詳細にも焦点をあてつつ報告がなされている。ほんのわずかな写本の断片や文字の跡から写本全体や語を復元するプロセスは、砂を積むような精緻で根気のいる仕事であり、この過程を経てようやく「読む」ことが可能なテキストがすがたを現すことを知れば、人文学の営為が、意味以前の痕跡から開始され、ようやく意味に至ることが理解されるだろう。ここで採り上げられているデジタル化画像の組み合わせによる写本断片の再構成は、他分野においても取り組まれてきたことであり、日本においても、正倉院文書の復原に関して一定の実績がある。こうした作業を行うにあたっての困難を解消するための仕組みも徐々に用意されつつある。すでに SAT-DB にも実装している IIF、とりわけ Image overlay の機能は、IIF 対応で画像が公開されている場合には有用である。撮影機器に関しては、デジタル撮影時に定規が映り込むようにしておくことで、実際のサイズを比較的確実に把握することができ、その定規を機械認識することで複数の画像を実際の縮尺にあわせて伸縮させる仕組みも開発されつつある。定規を映し込むことは手間のかかる作業にはなるものの、デジタル基盤上での共通の尺度としてその機能は予想以上に大きい。

chapter.05 青野道彦「諸版対照テキストと註釈対象語句索引の作成をどうすすめるか—*Samantapāsādikā* の研究基盤を整備する」では、上座部仏教における戒律聖典に対する註釈書である *Samantapāsādikā* の研究基盤を整備するための取り組みについて紹介している。すなわち、TEI ガイドラインに準拠しつつ諸版本の対照テキストを作成する一方で、註釈対象語句の索引としての構造に関しても TEI ガイドラインに準拠させるべく検討を行い、内容に即した深いマークアップの実践にまでつなげている。後者の構造化は、註釈文献と原典との関係を可視化する上で有益であるとする。研究者の理解を広く共有可能な仕方で記述する TEI の応用は、人文学の要請から構築され、今なお改良

が続けられている TEI の特質を如実に活かすものであり、そこからのさまざまな発展的な研究が期待されるところである。

chapter.06 李乃琦「一切経音義全文データベースの構築と研究」は、高麗蔵には収録されているものの大正蔵に未収録の玄応撰『一切経音義』の写本・木版本を対校すべく全文データベースを構築し、その活用と研究への取り組みを報告する。李氏がこれまでに確認した9種の一切経音義日本古写本、および高麗本を分析するに際して、それぞれの全文テキストの電子化による一切経音義全文データベースの構築が、複数の写本に対する個別分析のみならず、諸本の並行的比較において、極めて重要な役割を果たしたとする。データベース構築にあたり、一切経音義の構成をあらかじめ分析し、それに従って構造化を行うことで比較を容易にした点は注目に値する。ほかにも国内外に調査すべき古写本の存在が確認されており、今後のさらなる成果が期待される。

chapter.07 宮崎展昌「チベット語大蔵経データベースの利用および本邦に伝存する漢語大蔵経とその調査の重要性と可能性」は、チベット語大蔵経と漢文大蔵経をその歴史的経緯とともに紹介した上で、現存する資料の一覧を提示し、そのうちデジタル媒体として利用可能なものについて注記する。ことに、木版大蔵経の刊行史における『普超三昧経』の状況に着目しつつ、TEI ガイドラインの可能性を提起している。大正蔵が参照する資料が対象となる課題についての論攷でありであり、SAT-DB における TEI 対応においてことに示唆深い提言をしている。

chapter.08 石井清純「引用出典検索・読解とデジタル化―曹洞宗学におけるデジタルアーカイブの活用」は、曹洞宗教理研究の動向とそこにおけるテキストデータベースの役割と、画像データベースを用いた曹洞宗学の可能性について論じ、デジタル化されたテキストや画像の活用が宗学という日本の歴史に固有な知識伝承の内部においても有効であることを示し、さらにオンラインでの「正法眼蔵国際読書会」が国際共同研究にとって極めて効果的であることを紹介している。日本文化やその精神的背景への国際的な関心の広がりの中で、最先端のデジタルツールを用いて世界の諸地域で生まれつつある読者をより深い理解へと導く場の創出は、これからの仏教研究のみならず日本研究にとって示唆されるところが多いだろう。

次に「Ⅱ 研究基盤を作る」である。

chapter.09 菱輪顕量「中世の手書き写本の OCR 翻刻テスト報告」は、凸版印刷との協力のもと、タイトルの通り、手書き写本の OCR 翻刻への取り組みの報告をなしたものである。15 世紀末に成立した、大小さまざまな字体で独特の崩し字が多い『諸法実相伝抄』^{しよほうじっそうでんしやう}、崩し字の仮名で書かれた『仏国禅師法語』^{ぶつこくぜんじ げう}、13 世紀に東大寺尊勝院で活躍した宗性の写本、のそれぞれについての試行と、結果の把握、今後の可能性が述べられている。木版本を対象としたくずし字 OCR が成果を上げつつある中で、手書きの写本を対象とした実践報告には、技術的な可能性についての見通しを共有する上で有用である。SAT-DB では現在のところ扱う画像史料の多くは木版本であるが、日本撰述典籍を充実させていく過程ではこの試行は重要であり、その技術をめぐる状況には今後も注視していく必要がある。

chapter.10 王一凡「慧琳撰『一切経音義』の符号化をめぐる」では、大正蔵における UCS 未符号化文字を符号化提案するにあたっての現状と課題について報告し、特に題材として大正蔵における『一切経音義』『統一切経音義』を採り上げる。これらの音義書は、未符号化文字が非常に多いにも関わらず依拠できる資料が少ないために符号化提案が難しく、現在も慎重に続けられている検討内容が紹介されている。そして、SAT 研究会が取り組んでいる UCS 符号化提案の具体的な手続きとそれを取り巻く状況についてもやや詳しい説明が行われている。漢字の符号化について、学界として ISO の登録に関与しているのは SAT が唯一の団体である。これは漢字文化圏の人文学におけるまさに基盤の形成に関わるプロジェクトであり、今後、ほかの領域における同様の活動を支える貴重な知見となるだろう。

chapter.11 宮崎泉「電子テキストの有効利用に関する雑感—文献資料のモデル構築の可能性」では、電子テキストとはいかなるものか、という問いについて、仏教研究の観点から網羅的な検討がなされている。原典の言語として想定されるサンスクリット語のテキストから漢文やチベット語に翻訳されるという仏教テキストの歴史的経緯と、それに研究者がどう向き合ってきたかを概観し、これが電子テキストとしてはどのように展開し得るかという課題の検討である。具体的には、文字と書誌情報に着目し、文字コードの問題から文字情報

サービス環境 CHISE の設計思想を介してテキストの単位を包摂粒度として捉え直した上で、国際図書館連盟が定める IFLA LRM を踏まえつつ柔軟な書誌情報の在り方を提起する。SAT-DB では現在、大正蔵のテキストをもとにデータベースを構築しているが、そこから先に進む際には、本章で示される柔軟な捉え方を可能な限り反映する必要があるだろう。

chapter.12 苜米地等流「サンスクリット文献電子データについての雑想」は、サンスクリット文献の電子テキストの世界的な状況について概観した上で、その望ましい在り方について検討する。特に GRETIL が公開するテキストデータがかかえる問題に着目する点は、小野基氏の chapter.02 とあわせて読まれたい。他方、リソース提供のありかたについての問題提起は、DH における独創性や新規性のある研究開発と、利便性の高い情報の便利な利用者サービスという異なる次元での取り組みをいかに統合させるかという難しい課題である。現在生み出されつつあるデジタル学術空間にあっては、研究者個人の成果の集積がそのまま研究共同体全体の学術空間を形成するほどに研究基盤が成熟した段階にはない。この両者を、ある意味、試行錯誤によってつないでゆく時代がしばらく続くだろう。これは従来の人文学にその基盤が与えられて以降、多くの人文学研究者が見落としてきた、自身の学問の成立に関わる課題でもある。SAT-DB の開発は、こうした両立困難な問題を自覚しつつ、研究者個人のレベルでは対応がなしかねないだろう、後者の立場に立って進めてきた。一方で、データの提供が、開発者の自己満足に陥らないようにすることは、極めて重要な課題である。Digital Humanities 学会においても、少なからぬ研究がその域を出ていないように見受けられる。その意味でここは注目すべき chapter である。

chapter.13 落合俊典「蘇州西園寺蔵『大正新脩大蔵経既刊分一覽』（昭和五年四月現在）に見られる刊行予定書目」は、大正蔵日本撰述部の刊行開始当初の頃に企図されていたと思われる刊行予定書目が、蘇州市の西園寺の藏經樓に所蔵されているのを発見した貴重な報告を、第 57 巻以降における現在の大正蔵との対比も含めた一覽表の提供とともになしている。大正蔵の編纂について、これまで見えなかった過程が明らかにされたことは、大正蔵のデジタル版を提供する SAT-DB にとって、そのデータベースの意義を相対化し、日本の仏典をデジタル研究基盤として提供する、将来の展開を模索する上で、貴重な示唆で

ある。デジタル化以前の歴史を明らかにしてゆくことは、デジタル学術空間をより適切に構築するために、重要な試みである。

chapter.14 高橋晃一「研究者による情報発信としての「学術ウェブサイト」の評価の行方」は、Web を用いた学術情報発信の有効性を論じ、事例として、研究者自身による学術論文の公開とオープンアクセスの状況に触れた上で、研究支援ツールとしての Web サイトを挙げている。そして、後者についても有用性を高めるためには専門知が不可欠であることを確認し、その専門家としての営みを評価する必要性を説いている。デジタル学術空間の形成にとって、貢献者の評価のありかたを検討することは不可欠なことであり、「prologue 情報通信革命と人文学の課題」においても述べた通り、欧米においては学会が主体となって声明文がつぎつぎに公表されている。この議論はそこで重要な役割を果たすことになるだろう。