

横断型デジタル学術基盤を目指して —SAT2018 の構築を通じて—

永崎研宣^{†1} 下田正弘^{†2} A. Charles Muller^{†2} 蓑輪頭量^{†2}

概要: デジタル学術基盤がもたらし得る福音は、学術領域や研究機関を横断した統合的な研究環境を実現し得る点にある。情報技術の圧倒的な進展に伴う IIF の普及や TEI のグローバル化など、文化資料を学術的に扱うための技術的な環境も徐々に整いつつある。本発表では、現代的な情報環境を活かしたデジタル研究基盤の一事例として、SAT 大蔵経テキストデータベースの新版、SAT2018 を採り上げ、今後のデジタル学術基盤の可能性について検討する。

キーワード: IIF, Word2vec, TEI, 仏教学

Toward a Comprehensive Digital Research Environment: Through a Case Study of SAT2018

Kiyonori Nagasaki^{†1} Masahiro Shimoda^{†2} A. Charles Muller^{†2}
Kenryo Minowa^{†2}

Abstract: Digital Research Infrastructures will be able to provide comprehensive research environments traversing research fields and institutions. Emerging of IIF and globalization of TEI have been forming a technical environment to treat digital cultural resources for academic use. This presentation reports possibilities of digital research infrastructures through a use case of the latest version of the text database of SAT Daizōkyō Text Database (SAT2018).

Keywords: IIF, Word2vec, TEI, Buddhist Studies

1. はじめに

SAT 大蔵経テキストデータベース研究会（代表：下田正弘）（以下、SAT 研究会）は、1994 年のプロジェクト開始以来、大正新脩大蔵経を中心とした仏教学研究のデジタル化に取り組んできており¹、2008 年 4 月には全文検索サービスを中心とした Web サイト SAT 大蔵経テキストデータベース（以下、SAT DB）を公開した。その後、人文学におけるデジタル研究基盤のモデルを構築するという観点から、2012 年、2015 年に大幅な改訂とサービスの追加を行ってきた²。2018 年 3 月末には、全文検索 Web サイト公開後 10 年を迎え、これまでの基本機能を継承しつつ、現時点での人文系学術情報 Web サービスにおいて提供し得る機能を可能な限り組み込む形で新版（以下、SAT2018）をリリースした。

タイトルに掲げる「横断型」とは、SAT DB においては、全文デジタルテキスト化の対象となった大正新脩大蔵経が、すでに、領域横断的に多くの分野から、世界を横断して広く利用されてきたものであり、国内外の資料を横断的に収集して底本としたことなど、紙の時代からすでに実現されてきたことであり、そうした横断性は、SAT DB が公開された 2008 年からすでにデジタル媒体上でも実現されてい

たと言える。それでもなお本稿にて横断型としたのは、昨今のデジタル技術の動向にともない、新たな横断性が実現し得る環境が整いつつあるためである。本発表では、横断型デジタル学術基盤という観点を踏まえつつ、今回追加した新機能を中心として SAT2018 について報告する。

2. 実装の方針

SAT2018 の実装では、これまで基本としてきた下記の方針を引き続き可能な限り反映することを目指した。

- (1) 既存のオープンソースソフトウェアを組み合わせて可能な限りカスタマイズせずに利用する
- (2) データ形式に関しては既存の標準的な規格やガイドラインに可能な限り準拠する
- (3) 利用者講習会等を経て寄せられたフィードバックに対応する

まず、それぞれの項目に関する状況について以下に概観してみよう。

(1)に関しては、テキスト全文検索には Apache Solr 7.1.0、テキストデータの管理には PostgreSQL 10.3、IIF 画像配信には IIP Image Server 1.0、IIF 画像表示には Mirador2 を用

^{†1} 一般財団法人人文情報学研究所
International Institute for Digital Humanities
^{†2} 東京大学 人文社会系研究科

いている。そして、Web サーバには Linux (Redhat7), Apache2.4, PHP5.6, Python3.6 等, Web インターフェイス構築のために jQuery3, jQuery UI1.12, jQuery Mobile, Bootstrap3 を利用している。これらの多くは、ドキュメントが充実し、動作も比較的安定していることから、開発者側としても利用者側としてもそれほど導入・利用のハードルが高くないと言え、人文系データベース構築の際には有力な選択肢になり得るだろう。

(2)に関して、SAT2018 において明示的に採用したものについては、いくつかのレイヤーがあるため、少し整理してみたい。まず、(2a) 対象となる仏典テキストを扱う専門家コミュニティにおいて通用する標準が存在する。たとえば、統一的な経典名やそれを同定する経典番号、経典内の巻、頁、行といった単位である。次に、(2b)人文系分野を主に対象とする標準的なガイドラインとして Text Encoding Initiative ガイドライン(TEI-G)がある。また、(2c)Web 標準ではないがそれを志向する規格のような位置づけとして IIIF (International Image Interoperability Framework), (2d)コンピュータ全般に広く利用される規格として Unicode がある。さらにここでは、(2e) 著作物等について一般に広く用いられるクリエイティブ・コモンズ・ライセンスも挙げておきたい。

(3)については、2012 年の北海道大学を皮切りに、SAT の利用者に向けた講習会を国内外各地で開催してきており、そこでのフィードバックのうち SAT 研究会として今回対応可能な機能については実装を行った。ここでそのすべてを挙げることは困難だが、例を挙げるなら、(3a)検索条件を再開時に再利用できること、(3b)正規表現検索対応、(3c)同時に複数の経典を表示したり簡単に表示を切り替えたりできること、(3d)英語訳だけでなく日本語訳も表示できるようにすること、等である。また、利用者講習会では抽象的な問いや要望が出ることもあり、なかには(3e)人工知能が発達すると仏典や仏教の理解はどう変わっていくのか、というものもあった。

さて、以下の章では、ここまでみてきたトピックスのうちいくつかについて、それぞれに付した番号を参照しつつ報告していきたい。

3. オープンソースソフトウェアの活用(1)

これに関しては、ほとんどは標準的な使い方をしてに過ぎないので特筆すべきことはほとんどない。モバイル対応に関しては、Bootstrap3 及び jQuery Mobile の採用によって多くの機能が利用できるようになっている。Mirador がモバイル対応しているため、IIIF 対応画像に関しても問題なく利用可能である。このような様々なフリーソフトウェ

アはそれぞれに固有の背景を持っており、それらを横断して一つのデジタル学術基盤を形成しているとみるなら、これも「横断型」の一つの形と考えてもよいだろう。

なお、この文脈で特筆すべき点として、Apache Solr における検索対象語の総登場件数の計数がある。Apache Solr では検索対象語の関連度を測るために検索対象インデックスから検索語の数を数えていると思われるが、開発者である筆者の能力と与えられた時間では、その数値を取り出す方法を見つけることができなかった。そこで、今回の場合、検索対象の容量が限定されていることから、サーバサイドで用いている PHP で、ヒットしたドキュメントを読み込みその都度数を数えるという仕組みを用意することとした。ただし、一度にすべてを読み込んでしまうとサーバ側のメモリに容量が足りなくなる可能性があるため、一定数のヒットドキュメントごとに計数することを繰り返す仕組みを志向した。この計数を何件ごとに行うべきかを検討するために、いくつかの検索パターンについて実際に検索を行ってそれに要した時間を比較して決めることとして、以下のような結果を得た。具体的には、実運用に近い状態とするため、50, 100, 200, 500, 1000, 2000, 5000, 10000, 20000, 50000 件ごとに計数するアクセスをそれぞれ同時に 10 件実行してその平均値をとってみた。(図 1)

対象とした検索語 (ヒット数)

/世尊/ (110837 件)

/仏説|佛説|仏説|佛説/(17387 件)

/華嚴經|華嚴經|華嚴經|華嚴經/ (3986 件)

/照見五蘊皆空|照見五蘊皆空/ (42 件)

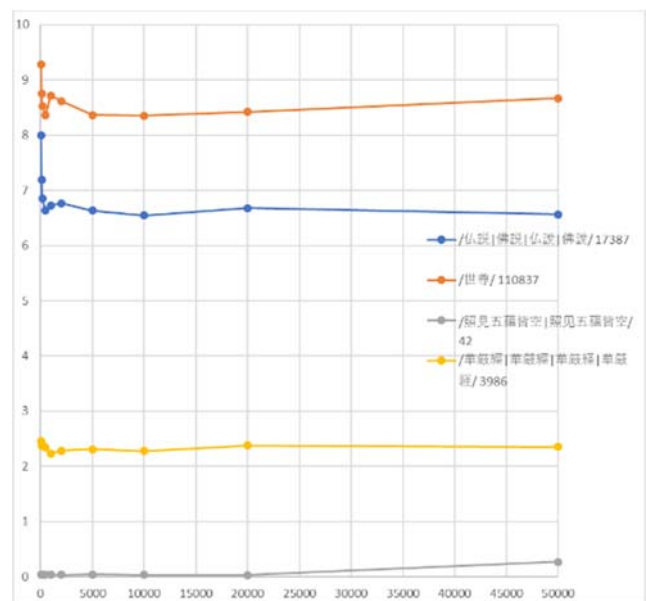


図 1 計数単位と検索に要した時間 (秒)

この結果はサーバの性能や各種ソフトウェアの設定^aに依

^a 本側の設定としては php.ini の memory_limit を 1024M としたこと以外はほぼデフォルト値のままである。サーバ側の設定の工夫についても今後の課

^a ここで使用している機材は Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz ×2 (計 32 コア)、DRAM64MB という構成である。この検索に関わるサー

存するところが大きいと思われるが、この環境では、10000件ごとに計数するパターンがどの検索文字列でも比較的速かったため、今回では10000件ごとに計数することとした。なお、SAT2018のWebインターフェイス上では、この検索対象単語の総数はAJAXを用いて非同期で問い合わせしているため、検索結果表示を長時間待たされるといった問題は生じない。

また、Apache Solrでは、正規表現検索が可能となっているが、これはインデックスにCJKのn-gramインデックスを用いようとする場合、n-gram設定の最大値が正規表現対象文字列の最大文字数ということになるようである。SAT2018では、8-gramまでの設定ができたが、9-gram以上ではインデックス作成が途中で止まってしまい、設定することができなかった。したがって、正規表現検索対応文字列は8文字までとなっている。ここでのインデックス用に作成したJSONファイルは590MBであり、これに対して最大8-gramのインデックスは9.9GBとなっている。この点については、今後、より長い検索文字列に対応できるように検討していきたい。また、正規表現検索対応させたインデックスの設定では近傍検索が動作しなかったため、同じデータを利用しつつ近傍検索用に設定したインデックスも別途作成した。そして、双方を効率的に利用できるようにするため、ユーザが検索した際に検索パラメータをみてどちらのインデックスを利用するか切り替える検索用ラッパーをサーバ側に用意した。

4. 標準的な規格やガイドラインに関して

4.1 専門家コミュニティで通用する標準的な事柄(2a)

大正新脩大藏経は収録典籍ごとに番号が付されており、その番号に対応する経典等を手がかりに様々な写本・刊本が同定されてきている。これは4世紀の道安の手になる『綜理衆経目録』に始まる中国語訳・中国語仏典の伝統に連なるものであり、大正新脩大藏経がとくに国際的に広く普及したためにその目録番号がそのままデジタルにおいても用いられるようになり、テキストの一部を参照する際にはその頁・行番号等も活用されている。同様のレイヤーと位置づけ得るものには、たとえば日本古典籍総合目録データベースにおける統一書名、ギリシャ・ラテンにおけるCTS (Canonical Text Services) などがあり、いずれも専門家の間で用いられ、研究基盤の構築に活用されている。SAT2018でもこれまでに引き続き様々な関連資料との対応付けを通じた研究基盤の形成にこの番号を用いている。番号そのものは専門家向けだが、結果として、他分野の研究者や初学者、一般利用者に対する利便性を高めることにも大きく貢献している。ここでも、「横断」が実現されていると言える

だろう。

4.2 人文系研究者のための標準的な事柄(2b)

SAT2018ではこの種の標準的なものとしてTEI-Gを利用しているが、ISAD(G)やCIDOC-CRMなどもこのレイヤーとして考えてよいだろう。このレイヤーは、一定の広さをもった専門家集団によって定められたルールであり、(2a)に比較すると、資料の外形的な性質や用途に基づいて対象範囲や手法を定め、ルールを設定していると捉えることもできる。SAT2018では、現代日本語訳においてTEI-Gを採用してマークアップをおこなった上で、大正新脩大藏経本文とは上述の目録番号・行番号等を用いて文章ごとにリンクを設定し、パラレルコーパスとしても活用できるようにしている。このリンクの現代日本語訳側は、<s>タグにxml:idを付与することで対応づけており、基本的なDOM操作によってインターフェイス構築を行っている。パラレルコーパスとしての検索も可能であり、古典中国語の仏典テキストで検索するとその検索語を含む文章と対応する現代日本語訳が表示されるようになっている。また、人物名や地名、会話における話者等に関してもそれぞれにTEIの対応するタグを付与している。それらのタグも含め、SAT2018上ではTEI/XMLファイルをjQueryでパース・変換して閲覧用インターフェイスに表示し提供している。このデジタルテキスト群はポーランドデジタルな現代日本語であり、日本語であってもTEI準拠には特に問題は生じなかった。なお、これらのファイルは仏教研究としても人文情報学としてもアウトリーチを志向するものであり、TEI/XMLとしてダウンロードしてCC BYライセンスで利用できるようになっている。これもまた、初心者から専門家までを横断するものであり、TEI/XMLを利用することで人文学他分野や他言語の資料との相互運用を実現し得るという点でも横断的であると言えるだろう。

4.3 エコシステムとしてのIIIFの導入(2c)

SAT2018における主要な取組みの一つはIIIFに関するものである。現在、すでに世界中の文化機関からIIIF対応画像が公開されるようになり、そこには時折アジアや東アジアの資料が含まれ、さらにその一部として仏典の画像も含まれるようになってきている。たとえば、海外においても、フランス国立図書館が提供するデジタル画像公開サービスGallicaでは、敦煌文書のコレクションとして3000点以上をIIIF対応で公開しており、ここには多くの仏典画像が含まれている。あるいは、早くからIIIFに取り組んでいるバイエルン州立図書館で公開している東アジアデジタルコレクションのIIIF対応画像にも仏典画像が含まれており、他にもハーバード大学・ケンブリッジ大学・ブリティッシュ

題としたい。

コロンビア大学など、仏典画像を含む IIF 対応画像コレクションを公開する海外機関は徐々に広がってきている。日本においては、SAT 研究会がデジタル化公開を担った東京大学図書館の IIF 対応画像が仏典画像の点数としては最も多く、1813 件、19 万枚超である一方、やはり海外と同様に、仏典公開を目的とせず、所蔵コレクションの IIF 対応画像公開を行うなかで仏典画像も含まれるという例が出てきている。すでに数万件の IIF 対応コンテンツを公開している国文学研究資料館だけでなく、京都大学図書館、九州大学図書館、島根大学図書館など、徐々に増加しつつある。IIF 登場以前からの問題として、こういった各地に分散する画像に利用者が効率的にたどり着く方法がなかなか見当たらないということが挙げられており、SAT 研究会としては、各地の仏典画像の URL を収集し大正新脩大藏経目録番号と対応づけることでその支援を行ってきた。しかし、リンクをクリックした先の画像はそれぞれに異なる使い勝手になっており、必ずしも閲覧しやすいものではなく、なかには書誌情報の頁には簡単にリンクできないが画像頁にリンクをするとその書誌情報を得ることはできない、というサイトもあり、利便性を高めるにあたっては大きな障壁となっていた。

IIF 対応画像においては、こうした問題は技術的にはほぼ解決している。しかし、運用面も含めて解決できているわけではない。そこで、SAT2018 では、これを解決するために IIF Manifest の効果的な共有を目的とする Web コラボレーションプラットフォーム IIF Manifest for Buddhist Studies (IIF-BS) を開発した。ここでは、編集協力者が IIF Manifest URI をフォームに記入して登録すると、その Manifest URI, Attribution, License, Title 等の基本情報が PostgreSQL のテーブルに取り込まれるとともに、IIF Manifest ファイル内のテキストデータを含む項目については Apache Solr に取り込まれ、全文検索ができるようになる。さらに、登録された Manifest URI に対して、編集協力者が大正新脩大藏経目録番号、巻番号、開始・終了頁行番号を記入できるようになっている (図 2)。

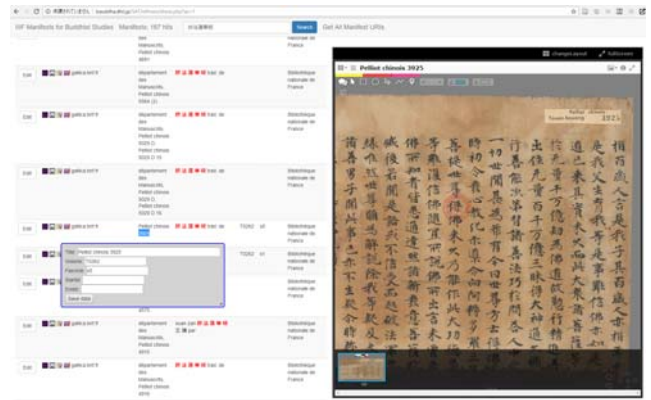


図 2 IIF-BS の作業画面

IIF-BS の情報は API のような形で外部サイトから利用でき、目録番号、巻番号等を URL で渡すとそれに対応する Manifest URI を JSON 形式で返戻するようになっている。SAT2018 では、各経典閲覧時にこの URL を読み込んで IIF アイコンを表示し、アイコンクリックで Mirador の画面を分割してその画像を表示するようにしている。(図 3)

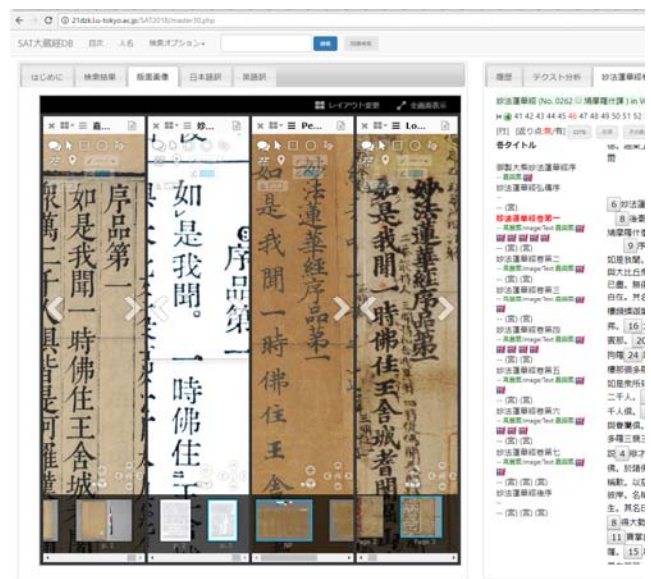


図 3 各地の IIF 対応画像を巻単位で IIF で提示し、妙法蓮華経の冒頭部分を並べて表示している例。

これによって、世界の IIF 対応画像を専門家コミュニティが自らのニーズにあわせて収集・編集し、そしてそれを誰もが機械的に再利用できるという状況が整備されたことになる。さらに言えば、各文化機関が自らのミッションとしてそれぞれに自らのコレクション等を Web コンテンツとして公開していくという各機関の固有の営みをとくに妨げることなく、一方で、専門家コミュニティが自らにとって

b たとえば、以下の URL で、妙法蓮華経 (T0262) の巻第 1 の IIF マニフェストを取得することができる。
<http://bauddha.dhii.jp/SAT/iifmani/show.php?m=getByCatNum&cnum=T0262&s>

crnm=s1

必要なコンテンツを自立的に収集・編集していくという自然な営みを行い、それによって各機関のコンテンツが横断的に活用されその価値がより高く認知されていく、という形になっている(図4)。ここでは、機関同士の壁の横断、機関と研究者コミュニティとの横断、といったことが実現され、一つのエコシステムが成立しつつあるとみることもできるだろう。

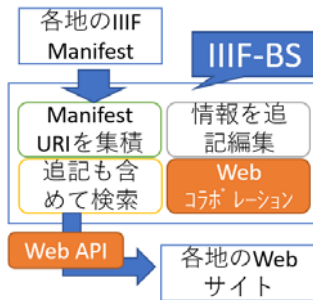


図4 IIF-BS を介した IIF コンテンツ利活用のイメージ

なお、IIF-BS では、SAT 研究会の認証データベースをそのまま利用しており、世界の 200 人程度の協力者が入力・編集できるようになっている。一方でこれはクラウドソーシングで行うことも可能であることから、Twitter アカウントで認証を行うものも用意し、2018 年 2 月 14 日に大阪大学で実施された日本文学研究者向け IIF ワークショップで実習用として利用に供し、十分に動作することを確認した。たとえば日本古典文学の IIF Manifest であれば、日本古典籍総合目録データベースの統一書名を利用することで IIF-BS のような機能を提供することは可能だろう。筆者らとしても、何らかのテーマを定めたシステムをクラウドソーシングで運用することも考慮したい。

また、IIF 関連としては、先行して公開している SAT 大正蔵画像 DB³の機能をほぼ取り込み、「画像検索」ボタンをクリックすると SAT 大正蔵画像 DB の Web API を通じて画像へのアノテーションを検索しその結果を表示するようになっている。そして、SAT2018 で導入した上述の機能をそのまま適用し、検索結果の各画像に付されたチェックボックスをクリックすると Mirador の画面を分割してその画像が拡大表示されるようになっている。アノテーションはタグ表示され、タグ中の検索要素をクリックするとその単語で検索できるという機能も SAT 大正蔵画像 DB と同様である。これは、技術的に横断しているというだけでなく、テキスト研究と画像研究という二つの領域を容易に横断できる環境が整えられたということもできるだろう。なお、IIF 対応コンテンツの場合、このようにして他のサイトにマージしたり新しいインターフェイスを付与したりすることが容易に可能であるという点も強調しておきたい。

4.4 Unicode に関して(2d)

SAT 研究会では、当初より UCS 未符号化の文字の扱いを改善するための取組みを続けてきており、2012 年、大正蔵に使用される UCS 未符号化文字の符号化提案を IRG (Ideographic Rapporteur Group) に対して行い⁴、2800 字超の漢字が 2017 年に Unicode10.0 として符号化された。SAT2018 では、これに準拠して Unicode への変換を行った。今後も引き続き、未提案文字の符号化に取り組んでいく予定である。Unicode における文字の符号化とその採用は、様々な課題はあるものの、現状では、他の電子テキストと SAT DB を横断的に利用する上で利便性を高めることに資するだろう。

4.5 ライセンスに関して(2e)

SAT 研究会では、東京大学総合図書館所蔵万暦版大藏経デジタル版の CC BY での公開⁵を皮切りに、それまでの学術利用限定・再配布不可というライセンスから、より再利用しやすいライセンスへの切り替えを検討してきた。とりわけ、欧米での東アジア・日本を対象としたデジタル・ヒューマニティーズの興隆は、再利用・再配布が容易なデータに対するニーズを飛躍的に高めている。SAT2018 では、それを受け、CC BY-SA でのデータ提供を行うこととした。大正新脩大藏経の IIF Presentation API のライセンス項目にはそれが記載されている。なお、CC ライセンスにおいては、対象コンテンツ中でパブリックドメインであると判断されるものには適用されないという注記があることから、部分的な利用については利用者の判断に委ねられることになる。これは、再利用可能なライセンスを付与することによる横断性を確保できるだけでなく、機械可読なライセンスを付与することにより、横断的な利用に対してより大きな可能性を開くことも期待される。

5. 利用者からのフィードバックに対して

利用者からのフィードバックは、開発者と利用者の垣根を越えた検討の場であるという点で横断的であると言える。SAT 研究会では、上述のとおり、2012 年より、世界各地の利用者講習会の開催という形でそれに取り組んできた。以下にその成果の一部を見てみよう。

5.1 (3a)検索条件を再開時に再利用できること

この機能に関しては、検索・閲覧操作の履歴を Web ブラウザの LocalStorage に保存するようにした上で、容易な操作でユーザが不要なもののみを削除できるようにした。SAT2018 にアクセスすると最初にこの履歴情報が表示されるようになっており、SAT2018 の新機能としては現在のところ好評を博している。

5.2 (3b)正規表現検索対応

正規表現検索対応については、上述のように Apache Solr の機能を利用することで対応を行ったが、今後の改良を通じてさらに利便性を高めていく必要があるだろう。

5.3 (3c)同時に複数の経典を表示したり簡単に表示を切り替えたりできること

これに関しては、以前は経典テキスト表示領域を1つにしていたのに対して、その表示領域を jQuery-tab で表示することにより、タブを増やして複数経典を簡単に切り替えられるようにした。また、タブ領域を左右二つとして、右領域での表示を基本としつつクリック一つで左領域にも表示できるようにした。

5.4 (3d)英語訳だけでなく日本語訳も表示できるようにすること

これに関しては、上述のように、現代日本語訳を表示し、大正新脩大藏経本文とリンクさせるとともに、パラレルコーパスとしても検索できるようにした。

5.5 (3e)人工知能が発達すると仏典や仏教の理解はどう変わっていくのか

この問題に対して直接の答えを出すことは困難だが、これに関連する近年の状況を体験的に考えてみるためのツールとして、Word2Vec によるテキスト分析機能を提供した。これは、大正蔵全体、印度・中国・日本撰述テキスト、各典籍の単位でそれぞれに Python3 向けに開発されたツールキットである gensim を用いてモデルを事前に作成しておき、ユーザは Web インターフェイスから任意の対象を選択すると Python の CGI でそのモデルから語彙とその関連語を取り出せるようにした。結果表示には Cytoscape.js⁶を用いて同時に2つまでグラフ表示できるようにして、ノードをクリックするとそのノードの単語で検索してさらにグラフが拡張していくという機能とした。また、ノードにカーソルをあてると、SAT DB が提供する Web API を介し、そのノードの単語で Digital Dictionary of Buddhism を検索して単語の意味を表示するようにしている⁶。結果として、利用者は、辞書で意味を参照しつつノードをクリックしながら関連語を探索できるようになっている。これを利用することで、たとえば以下のように、「空」という概念を説明するテキストに対する二つの注釈書の特徴が、Word2vec の機械学習によってどのような違いとして表現されるかを確認することができる (図5)。

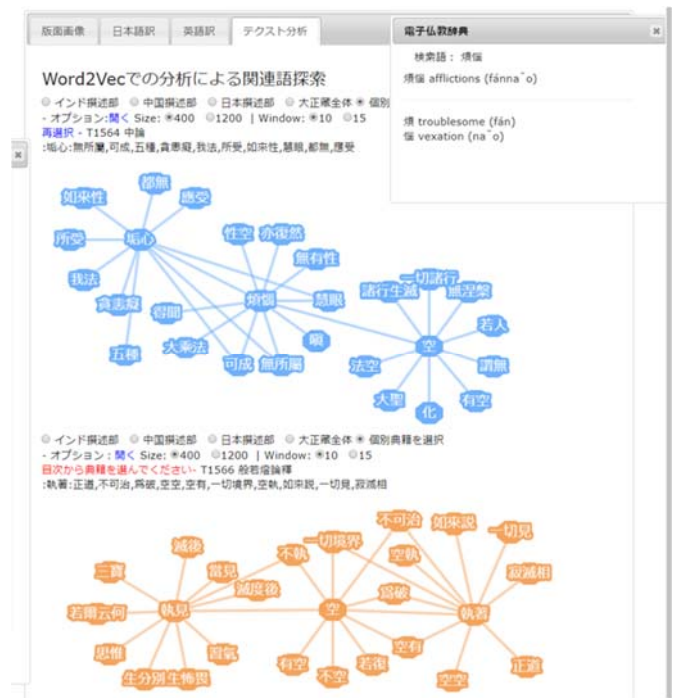


図5 『中論』の二つの注釈書に対する分析結果の対比

これは、仏教研究と情報学との横断する一つのソリューションであると同時に、今後の両者のさらなる横断的な関係の筋道を立てていくための一歩にもなるだろう。

6. おわりに

以上、横断性という観点を踏まえつつ、SAT2018 の新機能についての概要を説明してきた。SAT DB が当初より有してきた横断性は、IIIF への対応や Word2vec による分析機能の搭載を通じ、さらに強まることとなった。とりわけ、IIIF の潮流におけるデジタルコンテンツ公開機関から発された横断性への希求は、単に SAT2018 という専門サイトから横断するだけでなく、IIIF-BS という横断を促進するためのツールを介することで、より実用的かつ実践的な形で実現されることとなった。IIIF-BS の手法は、IIIF 対応コンテンツであれば特に分野を限ることなく様々な独自の文脈を付与して活用を図ることが可能である。今後は、単に SAT2018 の高度化を進めるだけでなく、IIIF-BS の可能性をさらに追求することで、横断的なデジタル学術基盤としての利便性を高めるとともに、その輪が各地でさらに横断的に幅広く形成され、仏教学のみならず、人文学全体をも裨益していくことを期待したい。

謝辞 SAT2018 は、国内外の多くの研究者の協力がなければ、完成に至ることはありませんでした。ことに現代日本語訳の作成は、公益財団法人全日本仏教会の財政支援に

c <http://js.cytoscape.org/>

よって実現し、また現代語作成の過程では、公益財団法人
仏教伝道協会から出版されている英訳大蔵經の成果を参照
することができました。本研究成果の一部は、JSPS 科研費
15H05725 の助成によって実現されたものです。

参考文献

- [1] 永崎研宣, 鈴木隆泰, 下田正弘 「大正新脩大蔵經テキストデ
ータベース構築のためのコラボレーションシステムの開発」
『情報処理学会研究報告』CH-70(2006年5月), pp. 33-40.
- [2] 永崎研宣, SAT 大蔵經テキストデータベース 人文学におけ
るオープンデータの活用に向けて, 情報管理, 58 卷 (2015) 6
号 p. 422-437, DOI: <https://doi.org/10.1241/johokanri.58.422>
- [3] 永崎研宣, 津田徹英, 下田正弘 「SAT 大正蔵画像 DB をめぐ
るコラボレーションの可能性」『情報処理学会研究報告』
2017-CH-113(8) (2017年1月), pp. 1-4.
- [4] 永崎研宣, 清水元広, 下田正弘 「UCS 符号化提案におけるデ
ジタルツールの活用—大正新脩大蔵經外字の符号化提案にあ
たって」『情報処理学会研究報告』2013-CH-97(4) (2013年1
月), pp. 1-6.
- [5] 永崎研宣 「大学図書館とデジタル人文学」『大学図書館研
究』(104) (2016年11月), pp. 1-10
- [6] A. Charles Muller, K. Nagasaki and Jean Soulat, “The XML-Based
DDB: The DDB Document Structure and the P5 Dictionary
Module; New Developments of DDB Interoperation and Access”,
Chung-Hwa Buddhist Journal, vol. 25, 2012, pp. 105-128.