

研究者から高校生までもターゲットとする
多機能仏典デジタルアーカイブ
SAT2018 がリリースされました

1. 発表者

下田正弘（東京大学大学院人文社会系研究科・次世代人文学開発センター人文情報学拠点長
インド哲学仏教学専門分野 教授）

2. 発表のポイント

次世代人文学のためのデジタル研究基盤のモデルとして開発された多機能仏典デジタルアーカイブ SAT2018 をリリースしました。

SAT2018 は、AI 技術を活用した大規模仏典テキスト探索機能等を実装することにより、精緻な専門研究の環境を整備すると同時に、平易な現代日本語訳仏典のオープンデータ公開をふくむ知識を一般社会に提供することで、人類文化に寄与する人文学的伝統知の社会全体の利用に向けた統合的知識環境を実現するものです。

3. 発表概要

SAT 大蔵経テキストデータベース研究会（代表：東京大学大学院人文社会系研究科下田正弘教授、以下、SAT 研究会）は、次世代人文学のモデル構築を目指す研究プロジェクトの一環として、新たなデジタルアーカイブ SAT2018 を公開しました。

SAT 研究会は、2008 年 4 月に最初の Web 版を公開して以来、2012 年、2015 年に大幅な改訂をおこない、近年では国内外から毎月 20-30 万件余のアクセス数を記録しています。この成果をさらに大きく更革する SAT2018 は、デジタル研究基盤として、大規模仏典集成「大正新脩大蔵経」の、AI 技術を活用した全文探索機能をはじめとする種々の研究支援機能を提供すると同時に、社会へのアウトリーチとして、中高生にも理解可能な現代日本語訳の仏典を「大正新脩大蔵経」の原文とリンクさせつつ、オープンライセンスで公開しました。

SAT2018 の完成によって、直接の専門となる仏教研究者にとって多機能で実用性の高い統合的研究環境が整備されるとともに、広く人文学研究者に対して専門知識デジタルアーカイブの先進的モデルが提示されます。加えて、この多機能仏典アーカイブは、専門的研究基盤を一般社会に開放する機能を備えており、人文学資料が広く人口に膾炙されることによって、人文学が一般社会に支持基盤を獲得するためのモデルとしての意義を有しています。

(SAT2018 の URL: <http://21dzk.l.u-tokyo.ac.jp/SAT2018/>)

4. 発表内容：

SAT 大蔵経テキストデータベース研究会（代表：東京大学大学院人文社会系研究科下田正弘教授、以下、SAT 研究会）は、1994 年、東アジア世界の仏典テキストの典拠として世界的に広く普及していた仏典の一大コーパスである「大正新脩大蔵経」（1 億字超）を、デジタル時代のテキストへと発展的に継承すべく設立されました。その後、2008 年 4 月に最初の Web 版を公開して以来、2012 年、2015 年に大幅な更新を行い、近年では国内外から毎月 20-30 万件余のアクセス数を記録しています。

SAT 研究会は、急速な発展をつづけるデジタル環境を人文学においていかに実現するかを課題とし、国内外の関連機関と広く連携しつつ、研究開発を進めてまいりました。この新たな環

境の実現のためには、研究の前提となる研究基盤を制度的側面から整備することも重要な課題となります。SAT 研究会は、仏典に出現する漢字外字（2017 年、約 2800 字種）や日本に固有の梵字異体字などの学術利用に必要な文字（6 文字）を Unicode に登録するなど、国際標準規格への制度的対応に関わる活動も持続してきました。

今回、公開しました SAT2018 は、これまでの活動成果全体を反映させることによって、資料となる仏典の専門的利用に関し研究支援を効果的に進めるデジタル研究基盤を提供することができました。具体的には、

- (a)自然言語処理を対象とした機械学習の一つである Word2Vec を用いた、「大正新脩大藏経」の関連語探索機能と専門用語辞書による探索支援機能、
- (b)Web 画像共有のためのデジタルアーカイブの最新国際規格 IIIF を用い仏典全文テキストデータと関連づけられた資料画像閲覧機能、
- (c)人文学テキスト資料の構造化ガイドライン TEI に準拠したテキストファイル及びそれを活用した機能、

これらの研究支援機能を提供しています。

さらに公益財団法人全日本仏教会の財政支援により、平易な現代日本語訳の仏典を作成してオープンライセンスで公開し、「大正新脩大藏経」の原文とリンクさせました。これにより、一般の読者が、現代日本語訳から入って 12、13 世紀の根拠資料に至る経路が Web 上に提供されました。人類が継承した伝統知を、著作権等の制度的制約に縛られることなく、一般社会の人々が自由に享受する環境を提供することは、研究者が果たすべき新たな社会的使命として、国際的に要請が高まっています。SAT2018 は、この要請にいち早く応えた事例として、その社会的意義が注目されます。

SAT2018 の公開を通して、種々の最新技術や規格を採用したシステムが日本の人文学において稼働し、実体験されはじめることによって、デジタル時代の知の共有についての議論が日本に根づくとともに、仏典テキストに代表される人文学資料が広く人口に膾炙され、人類文化に関わる研究への理解が深まってゆくと期待されます。

今後、SAT 研究会としては、この仏典の現代語訳を拡充するとともに、デジタル研究基盤としての機能をいっそう高めるべく、関連する最新技術の調査分析と適切な導入を継続してゆく所存です。

(SAT2018 URL: <http://21dzk.l.u-tokyo.ac.jp/SAT2018/>)

5. 問い合わせ先：

大学院人文社会系研究科

教授 下田 正弘（しもだ まさひろ）

Tel： 03-5841-3757

6. 添付資料：

The screenshot shows a web application for Buddhist terminology. The left sidebar contains a search for 'ほう-くう【法空】' (Hō-kū【法空】) with a definition and a list of related terms. The main area displays 'Word2Vecでの分析による関連語探索' (Word2Vec analysis-based related word search) for '法空' and '探水'. Two network diagrams show semantic relationships between related terms.

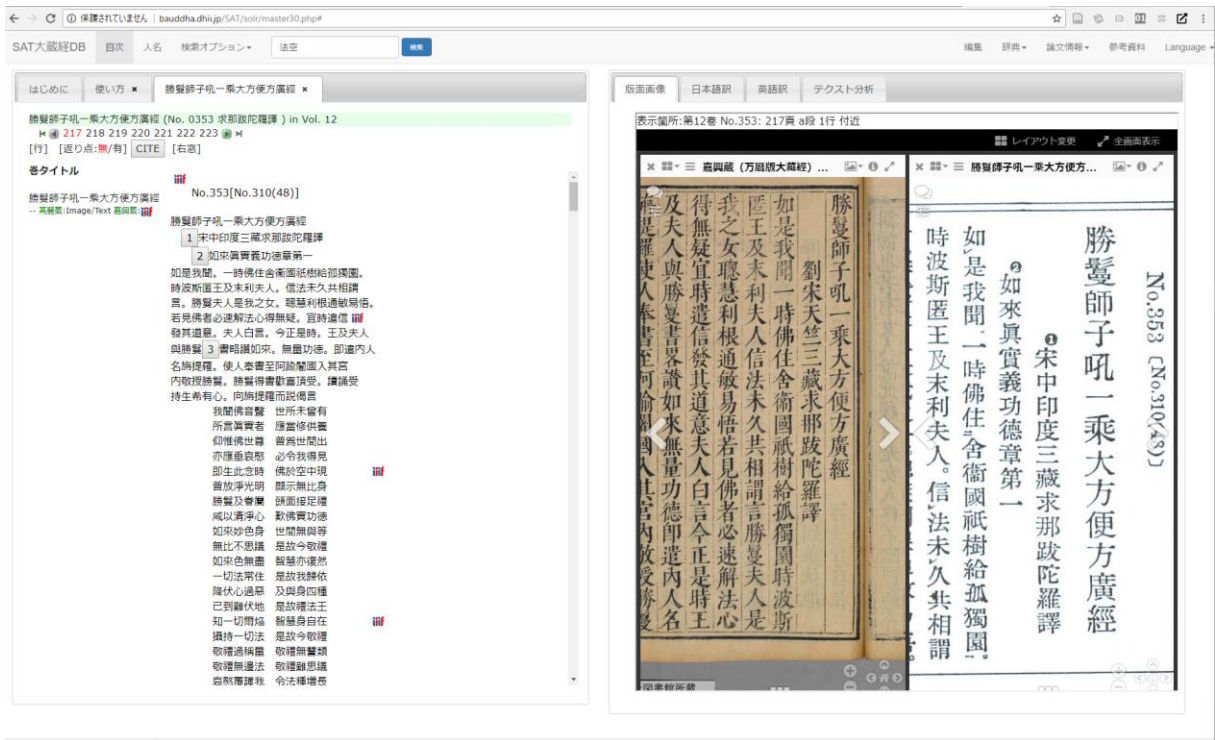
法空 (Hō-kū) Network Diagram:

- Central node: 法空
- Related terms: 自空, 行空, 空義, 諸法性, 空空, 無所有, 空無, 無倒性, 道妄情有理無, 周遍計度, 妄所執, 緣力所生, 依緣事, 依識變損滅執, 無相, 空即是色, 諸法空, 無我, 衆生空, 色空, 我空

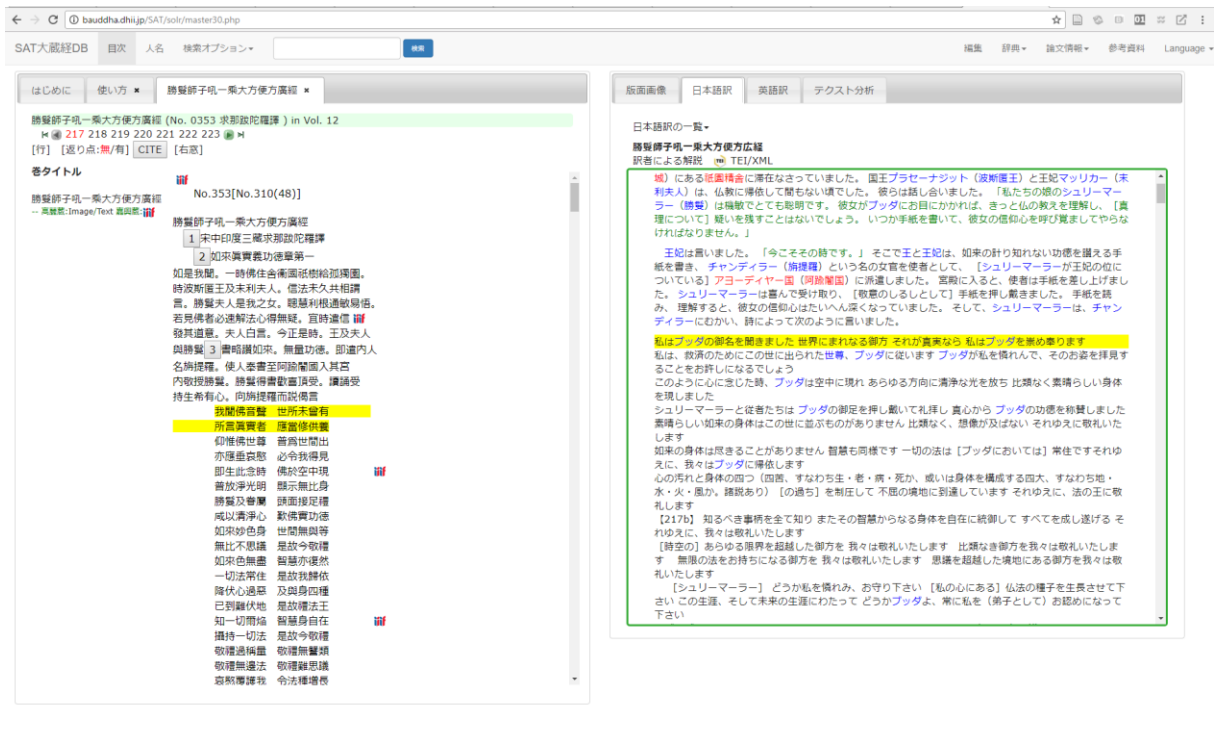
探水 (Tansui) Network Diagram:

- Central node: 探水
- Related terms: 団, 木蛇, 誰, 裏, 蕭寺, 昇, 褰衣, 心友, 老參, 澗, 野盤信, 雜, 他郷, 雞聲, 風, 落, 空, 嶺, 雲, 裡, 續, 沒交涉, 暗昏昏地, 破僧, 教

(AI 技術を活用した大正新脩大藏經探索機能と専門用語辞書による支援機能)



(国際的な Web 画像共有の枠組み IIF を用いたテキストと二つの版面画像の対比)



(現代語日本語訳仏典と大正新脩大藏經との対照閲覧機能)