

# 図書館における木版本のデジタル化と利活用の可能性 - IIF と TEI を用いた『水滸伝』諸版本のデジタル化を通じて -

上原究一\* 永井正勝\* 中村覚\* 中尾道子\* 近藤隼人\* 荒木達雄 袁輪顕量\*

\* 東京大学

本発表では、東京大学附属図書館アジア研究図書館上廣倫理財団寄付研究部門（Uehiro Project for the Asian Research Library）において新たに作業を開始した『水滸伝』諸版本のデジタル化及びその研究について、人文学の立場から研究上の課題を説明した後、情報学の立場からそれを解決するシステムの可能性を示し、デジタル化の先に広がる利活用の可能性についての提案を行う。

## Possibility of Digitalization and Utilization of Block Books in Research Libraries - Case of Digitization of “Shuihuzhuan” various editions using IIF and TEI

Kyuichi Uehara\*, Masakatsu Nagai\*, Satoru Nakamura\*, Michiko Nakao\*, Hayato Kondo\*, Tatsuo Araki, Kenryo Minowa\*

\* The University of Tokyo

In this presentation, after explaining research subjects from the standpoint of humanities about digitization and its research of the “Shuihuzhuan”(a.k.a “Water Margin”) various editions which started at the Uehiro Project for the Asian Research Library, The University of Tokyo Library System, will show the possibility of a system to solve the subjects from the standpoint of informatics, and propose the possibility of utilization spreading beyond the digitization.

### 1. はじめに

本発表は、東京大学で進行中の2つのプロジェクトの研究成果をもとに、木版本のデジタル化によって広がる利活用の可能性について提言を行うものである。

本研究に関与する1つ目のプロジェクトは、2016年度に開始された「東京大学デジタルアーカイブズ構築事業[1]」である。本事業は学内の学術資産の公開と利活用の支援を目的として立ち上げられ、附属図書館長、博物館長、文書館長等によって構成される委員会が実施している。この取り組みの中で、IIFやTEI等の国際標準規格に準拠して画像データ等を公開する「東京大学学術資産等アーカイブズ共用サーバ」を運用しており、本発表で対象とする附属図書館所蔵の『水滸伝』についても、IIFデータとして公開する予定である。しかしながら、本事業はあくまでもアーカイブ作成支援に特化した組織であり、資料に対するメタデータ付与は行っていない。この状況が必要となるのが、資料に熟達した研究者であり、資料のメタデータに始まる浅い記述から、本文のテキストデータの付与などに至る深い記述まで、学術的に求められる情報を研究者の立場から付与する必要がある。

この点において研究協力をしているのが、2つ目のプロジェクトとしての「東京大学附属図書館アジア研究図書館上廣倫理財団寄付研究部門（Uehiro Project for the Asian Research Library）以下U-PARLと呼ぶ」である。U-PARLは2014年に附属図書館に設置された研究部門である。U-PARLではこれまで東京大学総合図書館所蔵ならびにアジア研究図書館所蔵となる漢籍・碑帖拓本資料の高精細画像を作成し、現在、20点の資料をFlickr上で公開している。U-PARLは図書館所属の研究部門であるため、第一に書誌情報などの浅い記述を目指して、画像を公開した次第である。さらに、資料に対する深い記述として本文のテキストデータの付与などの作業を現在実施中である（その成果の一部は[2]で紹介した通りである）。

また本年度より、U-PARLで作成した画像・浅い記述・深い記述を東京大学デジタルアーカイブズ構築事業のもとでIIFとして統合し、TEIに準拠してマークアップしたテキストデータなど、研究成果に基づくデータ提供をすべく、作業を開始した。そこで本発表では、新たに作業を開始した『水滸伝』諸版本のデジタル化について、人文学の立場から研究上の課題を説明した後、情報学の立場からそれを解決するシステムの可能性を示し、デジタル化の先に広がる利活用の可能性についての提案を行う。

## 2. 『水滸伝』 版本研究に必要なこと

### 2.1 白話小説における版本研究の意義

今回そうしたデジタル化研究の対象として、中国明代の白話小説（口語体で書かれた小説）『水滸伝』を選んだが、以下にその理由を述べる。

白話小説は、16世紀半ば以降の木版印刷による商業出版の活発化に伴って隆盛を極めた娯楽的な文学ジャンルである。当時においては売れ筋の作品は様々な地域の多くの出版業者から同時並行的に刊行されるのが常であったが、その際には付加価値を加えるために増補・改編されたり、反対に出版コスト削減のために短縮・簡略化されたりと、商業的な思惑によってテキストが改変されることが多く、挿絵や附録といったおまげが付くことも珍しくなかった。このような性格を持つ白話小説の研究に当たっては、いかなる場合にも常に底本とすべき絶対的なテキストがあるというわけには行かず、研究の目的によって選ぶべき底本が変わることになる。例えば、作品を集大成した人物の創作（編集）意図を探ろうとするならば、なるべく商業的意図での改変が加わっていないテキストを用いるべきだし、ある作品のある時期における受容のされ方を調べるならば、その頃普及していたテキストに拠らねばならない。言わば、全てのエディションが各々独自の意味を持っており、各自の性格を把握する作業を経なくては研究が始まらないのである[3]。

また、木材である版木の水分の含有量の変化に伴う伸び縮みや、印刷を重ねることによる版木の損傷や、埋木改刻による意図的な文字の改変などによって、同じ版木から刷られた本でも印刷時期によって版面の様相が変化するという木版本の性質も相俟って、ある白話小説作品の研究のためには、理想的には現存するありとあらゆる本を調査することが望ましいということになってしまふ。しかし、世界各地に散らばる膨大な数の本を悉皆調査することは、もちろん現実的ではない。従来は重要と思われるエディションを影印出版することで研究の便を図っていた訳であるが、そもそも重要なエディションどうかを判断するためには、まず研究者による調査を経なければならない。よって、研究上大きな意義を持つエディションが人知れずどこかに眠っている事例は今なお少なくないと予想される。そこで所蔵機関によるデジタル化公開の意義が大きいのである。

数ある作品の中から特に『水滸伝』を選んだのは、白話小説の中でも最初期の作品にして最高峰の傑作の一つでもあることもさりながら、東京大学に属する3つの図書館・室にそれぞれ複数の稀観版本が所蔵されていることが決め手となった。附属図書館の研究部門としての、部局の枠を超えた研究の取り組みのモデルケースにもなると思えた次第である。現時点では、手始めにそのうち総合図書館所蔵の3種の木版本のデジタル化を行っている。なお、このうち2種は木版印刷によ

る挿絵を伴うため、テキストと挿絵の関連や、東アジア美術史の面からも研究対象となり得る資料である。

### 2.2 白話小説デジタル化研究の現在

白話小説のデジタル化研究においては、2001年からほぼ毎年開かれている国際学会「中国古典小説、戏曲文献暨数字化国際研讨会（中国古典小説と戯曲の文献学的並びにデジタル化研究国際学会）」を主宰している中国・首都師範大学の周文業教授が1990年代末から開発を始めて改良を重ねて続けている簡体字中国語版 Windows 専用のアプリケーション（通称：中国古典小説版本比較プログラム）があり、この分野の研究者には広く知られて様々に活用されている。

このアプリケーションには『水滸伝』『三国志演義』『西遊記』『金瓶梅』『紅樓夢』という明・清代の白話小説を代表する5作品それぞれの主要版本（基本的には影印本が出版済みのもの）の全ページの画像データと、それに基づいて作成された全文のテキストデータが収録されており、異なる版本間のテキストデータを比較してその相違点を見易い形で示す「分窗口対比（ウィンドウに分けての比較）」（図1）や「逐行対比（行を区切っての比較）」（図2）、ある版本のテキストデータと原本の画像データとを容易に対照する「图文対照（画像とテキストの対照）」（図3）、ある版本において近接する位置に現れる同じ語句の間に挟まれた文章を後継版本が脱文にしてしまう「同詞脱文（日本の書誌学用語では「目移り」と呼ばれている現象。版本系統を探る上での指標として非常に有用な情報となる）」をテキストデータから自動的に探し出す機能ほか、多彩な機能が2008年までに揃っている[4]。

図1では、2ないし3種類の版本のテキストデータの改行を無視して窓ごとに表示して比較し、それぞれの窓の中で同じ文字が同じ位置に来るように表示する機能を有している。版本間の大きな異同の有無が視覚的に把握しやすいが、数文字単位の細かい異同の分析にはやや不向きである。

図2は6種の版本による「逐行対比」を行った例を示す。任意の数の版本のテキストデータの改行を無視して指定した字数で1行ずつ縦に並べて表示して比較し、それぞれの行の中で同じ文字が同じ位置に来るように表示する機能を有する。大きな異同の把握には向かないが、細かい異同を視覚的に把握しやすい。

図3は「图文対照」の例を示す。左の窓に目次、中央の窓に版本の画像データ、右の窓にページごとに区切ったテキストデータを表示し、3つの窓が連動する機能を有する。



図 1 3種の版本による「分窗口对比」の例  
Figure 1 Example of interface to display texts from three editions in parallel.

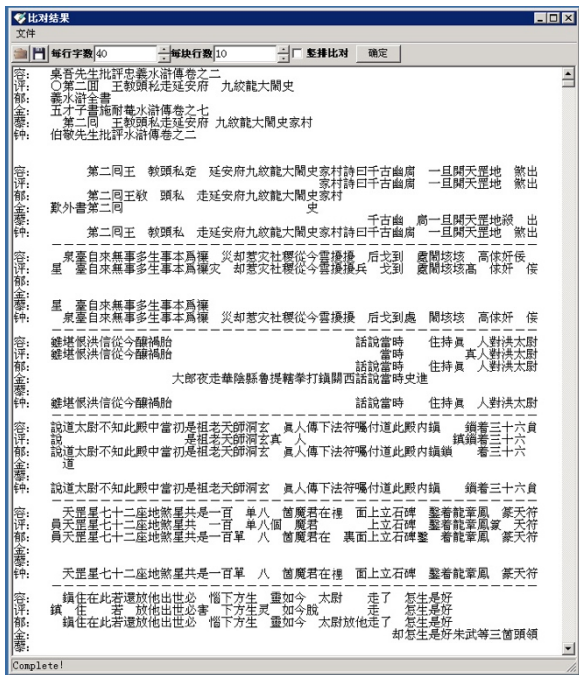


図 2 6種の版本による「逐行对比」を行った例  
Figure 2 Example of interface to compare texts by line from six editions in parallel.

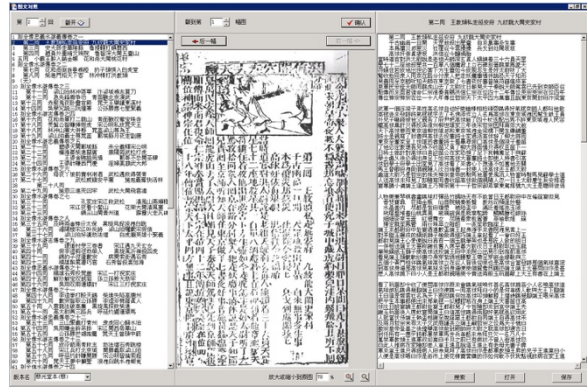


図 3 「图文对照」の例

Figure 3 Example to display text data and image data in parallel.

収録されているテキストデータは原本のチャプターに従ってファイルを分けており、必ず原本と同じ位置で改行し、原本のページの変わり目ごとに空行を入れてある。白話小説に多用されている俗字や異体字については、それ自体が研究の対象となるような重要な情報であるため、Unicode 拡張 B までの範囲で可能な限り忠実に入力する方針を採っている。なお、データが収録されていない版本についても、画像データとテキストデータを揃えることさえできれば、ユーザー自身の手で容易に追加が可能である。

このアプリケーションの抱える課題としては、中国人研究者が個人的に開発したものであるために、基本的には簡体字中国語版 Windows でしか動かせないという制約があることをまず挙げなければならない（それ以外の言語版の Windows でも、システムロケールを簡体字中国語に設定した状態であれば一応動かすことはできるのだが、それでは他のアプリケーションの利用に不便をきたしてしまう）。

また、画像データからテキストデータを産出する工程は基本的にこのような資料を扱うのに適した OCR 技術を持つ業者に外注しているそうなのだが、複雑な異体字の処理を多く行わねばならないこともあって、テキストデータには誤入力が散見されるという問題もある。このため、自分の目で画像と見比べてテキストデータを補正しながら使っているという研究者も少なくない。むしろ最終的には研究者自身の目で原資料を確認すべきであることは言うまでもなく、そのためにこそ「图文对照」の機能が用意されているのではあるが、「分窗口对比」や「逐行对比」の機能をより有効に活用するためにも、テキストデータの精度を高めることは重要な課題であると言える。

本研究ではこの先行事例を参考にしつつ、TEI や IIIF などの国際標準規格に準拠することにより、汎用的に利用可能な研究資源の構築を目指す。さらに、上記アプリケーションの備える機能のオンラインでのデジタル化公開への応用や、まだ実

現されていない機能の追及を目指しており、本稿ではその取り組みの一端を紹介する。具体的には、「图文対照」に相当する機能は、例えば Nicholas ら[5]による先行事例において、TEI 準拠のテキストデータと IIIF 準拠の画像データを関連づけて閲覧可能なビューアが開発されている。また「分窓口比対」についても、Versioning Machine 5.0[6]等、既に類似の機能が存在しているため、その活用が見込める。

テキストデータの作成に関しては、このアプリケーションに収録されているものを含めた既存のものを参考にしつつも、様々な分野の専門家が所属する研究部門によるプロジェクトであることの強みを活かして、白話小説を専門とする研究者による全面的な確認作業を行って精度の向上を図ることとした。

一方で、実現の目途が立っていない機能もある。木版本の研究においては、一見そっくりな2つの本が同じ版木で刷られたものか、それとも異なる版木で刷られたものなのかを見分ける作業や、

(同じ版木か別の版木かに関わらず) 両者の間に文字の異同があるか否かを判別する作業が非常に重要となる。例えば、図4の左図と右図は同じ版木で刷られながら下方中央に記された出版業者の名だけが差し替えられているという事例であり、左図が元の形で右図が改変後の形なのだが、こうした判断は専ら熟達した研究者の職人的な観察眼に頼るしかないのが現状である。それを自動的に行うところまでは行かずとも、デジタル化技術を用いて人の目による判断を補助する程度の機能が実現されるだけでも研究の飛躍的な効率化が見込まれる。例えば、北本ら[7]は差読 (different reading) により、ほぼ同一の画像に埋もれた微妙な差異を強調する手法を提案している。人文学の方面から真に求められるこうした機能の実現に向け、各種情報技術の活用が期待される。



図4 左:『新刻全像忠義水滸傳』東京大学総合図書館蔵本, 右:『新刻全像忠義水滸傳』ベルリン国立図書館蔵本

Figure 4 Left: “Xinke Quanxiang Zhongyi Shuihuzhuan” (held by General Library, The University of Tokyo Library System), Right: “Xinke Quanxiang Zhongyi Shuihuzhuan” (held by Staatsbibliothek zu Berlin).

## 2.3 白話小説の挿絵研究の現在

2.1 で述べた通り、主に営利目的で出版される娯楽作品であった白話小説の版本には、挿絵が付いていることが多い。その挿絵の形式や画風には、出版された地域や時期によって様々なタイプがあった。形式については、早くも宋代から商業出版が栄えていた地域である福建の建陽では上図下文と呼ばれる全ページの上部に挿絵を配する元代から流行していた形式が明代を通して主流であったのに対して、明代後期になって商業出版が本格化してくる南京や杭州や蘇州など江南地域では上図下文形式は稀で、見開きの片面または見開き全体をいっぱいに使った挿絵が多く見られる。画風の面では、万暦 (1573-1620) 前期には陰刻 (黒ベタ) を多用して画面にメリハリを付けつつ人物を大きく描く金陵派と呼ばれる画風が南京を中心に流行し、万暦後期以降は主に徽州出身の画工の手になる瓜実顔を特徴とした繊細で精緻な徽派と呼ばれる画風が隆盛を極めた[8]。

金陵派に属する画工一族である上元王氏は万暦 20 年 (1592) 前後に南京で出版された多くの白話小説の見開き形式の挿絵を務めているが、彼らが複数の作品をまたいで定番の構図の使いまわしを頻繁に行っていたこと[9]や、1つの版本の中で挿絵によって同じキャラクターの服装が本文における描写に忠実だったり相違していたりすることから署名を残す一人だけではなく複数の人物で作画を分担していたと考えられること[10]などが近年の研究で指摘されている。また、万暦 30 年代後半以降の蘇州や杭州では、万暦 20 年前後の上元王氏の見開き全面を使った挿絵の構図を見開き片面の形式に落とし込んだ徽派の

画風の挿絵を持つ版本が多いことも知られている[9][11].

このような研究成果は基本的に専ら人間の目で見比べることによって出されているものであるが、研究者が挿絵の画像に様々なタグ付けをして検索を可能にすることで様々な版本の中から類似要素を持つ挿絵を見つけやすくなることができれば、そうした人間の目による研究をサポートすることが期待されるであろう。作品をまたいだ影響関係が頻繁に認められるため、こうしたタグ付けは『水滸伝』だけで実施するよりも、多くの作品について網羅的に行うことが望ましい。それどころか、白話小説という枠にとらわれることなく、ジャンルを問わない幅広い明清刊本の挿絵を対象とすることによってこそ、より価値のある発見を導ける可能性が大である。元よりすぐに実現できる訳ではないが、将来的な展望として記しておく次第である。

また、こうした中国版本の挿絵のデジタル化研究は、東アジア美術史研究にも敷衍することができる。中国の白話小説の版本は中国国内のみならず、朝鮮や日本にも伝来・流入し、広く東アジア各地域で読み継がれてきた。とくに『水滸伝』をはじめとする挿絵を伴う中国版本は朝鮮や日本の絵画にも影響を与えており、中国版本から取材し、挿絵の図像を採用した作画が確実に行われていたことが知られている。諸版本の同一図題の画像を抽出したり、キャラクターやアトリビュート（持物）ごとに図像を採集し、様々な切り口でタグ付けを行って再編可能な状態で提供することで、画家たちが、いつの時代の、どの地域の、どのバージョンの刊本を見ていたのかという推測も可能となるであろう。また、同一画面上での図像の表示が可能になれば、表現の比較のみならず、当地で現在は失われていたり、改変や省略が行われている図像の考証にも有益となる。日本や朝鮮の絵画の典拠となった中国版本の挿絵のデジタル化を行うことは、書誌学研究のみならず、美術史をはじめとする隣接諸分野の研究にも資するところは大きい。

### 3. IIIF と TEI を用いた研究資源の構築

#### 3.1 はじめに

ここでは2で整理した『水滸伝』版本研究に求められる要件に基づき、この要件を満たす研究資源の構築に関する取り組みについて述べる。基本的には、国際標準規格である IIIF や TEI に準拠し、第三者が多様な目的に応じて再利用可能なデータの作成を行う。さらに、既存ソフトウェアの利用や、新規アプリケーションの開発を通じ、作成データが『水滸伝』版本研究の利用に資することを確認する。

#### 3.2 テキストをベースとした研究資源構築

2.2 では、『水滸伝』の本文データを対象とした研究において、異なる版本間の本文の異同の有無をテキストデータによって視覚的に認識しやすい形で提示する機能や、テキストデータとその元となった原本の画像データとを並べて確認できるようにする機能が求められることを述べた。このようなテキストに基づく版本比較分析を支援する研究資源の構築を行う。具体的には「TEIを用いた版本比較支援」「IIIF と TEI を用いた本文テキストと画像の閲覧支援」を目的としたデータの作成を行う。

##### 3.2.1 版本比較支援に向けた研究資源の構築

TEI では、版本の比較を支援するためのタグセットが提供されている。また、先述した Versioning Machine 5.0 や TEI Critical Apparatus Toolbox[12]など、表示・分析支援ツールも開発されている。したがって、ここでは TEI が提供する Critical Apparatus (校本) を記述するためのタグセットを用いて、版本の異同を記述し、版本の比較研究に資するデータを作成することを目的とする。

具体的には、東京大学附属図書館所蔵の『水滸伝』2種を対象として、当該タグセットによるマークアップを試みる。app (apparatus entry) タグを用いて異同箇所をマークアップし、rdg (reading) タグを用いてその読みを記述した。

以上で作成した TEI データを校本表示ツールで表示した例を図5に示す。本原稿執筆時点においては、対象資料の一部テキストのマークアップに留まるが、異なる版本間における本文の異同を確認することが可能となった。



図 5 本文テキストデータの比較例(上: Versioning Machine 5.0, 下: TEI Critical Apparatus Toolbox)  
Figure 5 Example of text comparison.

### 3.2.2 テキストデータと画像の関連付け

ここでは、テキストデータとその元になった原本の画像の閲覧支援に向け、TEI 準拠のテキストデータと IIIF 準拠の画像を関連づけたデータ作成を行う。TEI と IIIF の関連づけについては、先述した Nicholas らによる先行事例が存在し、マークアップ手法の提案や『延喜式』を対象としたビューアの開発がなされている。本研究では、この先行事例をベースとし、『延喜式』以外のテキストデータにも当該マークアップを適用可能とする拡張を行い、『水滸伝』への適用を試みる。

具体的なマークアップ手法の例を図 6 に示す。まず、対象とする画像群に関する IIIF マニフェストの URI を `surfaceGrp` タグの `facs` 属性に与える。また、各画像に関して、`Image URL` と `Canvas ID` を、それぞれ `graphic` タグの `url` 属性、`n` 属性に与える。これにより、`zone` タグによって範囲選択されたテキストデータを、関連づけた IIIF マニフェストの該当箇所に関連づけることが可能となる。本研究では、『水滸伝』の行単位で、`zone` タグとの関連付けを行なった。



図 6 TEI と IIIF を関連付けるマークアップの例  
Figure 6 Markup example to integrate IIIF into TEI.

以上で作成したデータを、既存研究をベースとして改良したビューアで表示した例を図 7 に示す。図左部にテキストデータ、図右部に画像が表示される。テキストデータ部のカメラアイコンをクリックすると、当該テキストが出現する画像箇所にフォーカスする。これにより、本原稿執筆時点においては、対象資料の一部テキストのマークアップに留まるが、本文のテキストデータとその元となった原本の画像を同時に閲覧することを可能とした。

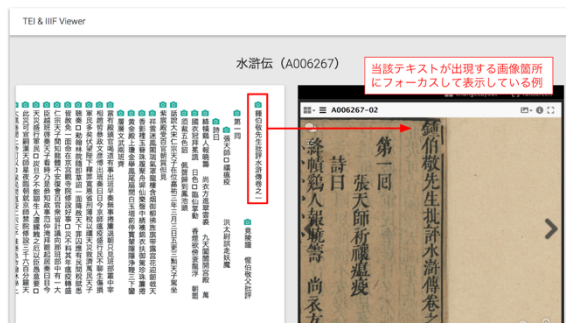


図 7 TEI と IIIF の関連づけによるテキストデータと画像データの並行表示例  
Figure 7 Example to display TEI based text data and IIIF based image data.

なお、本ビューアは図 7 に示した方式でマークアップした任意の TEI データを読み込むことができる。また、ブラウザ上での TEI データの表示については、JavaScript ライブラリ「CETELcean[13]」を用いている。

### 3.3 画像をベースとした研究資源構築

次に、2.3 で挙げた、白話小説版の挿絵を用いた研究への活用に向け、挿絵に付された題目(図題)や、挿絵に含まれる人物や道具のデータベース化を検討している。これには、CODH (人文学オープンデータ共同利用センター) が開発している IIIF Curation Viewer[14]を利用する。本ビューアは、IIIF 準拠画像に対する部分矩形領域の切り出し、利用者独自のコレクションの構築を可

能とする。本ソフトウェアの活用例として、鈴木らによる美術作品に出現する顔の部分を取り取って集め、それを美術史研究（特に様式研究）に活用するプロジェクト「顔貌コレクション」がある[15]。本事例を参考として、白話小説の挿絵に関するキュレーションリストの作成を目指し、まずは今回デジタル化の対象とする3バージョンから手を付ける。これらのうち1つは図を欠いているが、1つは上図下文形式のもの、1つは見開きの片面に複数の場面の図を押し込める形式のものである。

具体的には、まずは後者の図について、1つ1つの場面の図を切り出しながら、そこに原本に書かれている図題のメタデータを付与する作業を試み始めたところである。IIIF Curation Finder[16]で表示した例を図8に示す。IIIF Curation Viewerを用いて、挿絵に関するキュレーションのリストを作成し、さらに各々の図題名や描かれている人物をメタデータとして登録している例を示す。

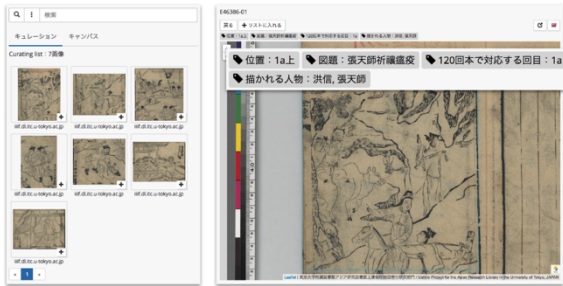


図8 挿絵に関するキュレーションリスト  
Figure 8 Curation list of illustrations.

### 3.4 作業上の課題

ここで挙げた利用例は、いずれも新規のものではない。重要な点としては、専門家による知識をIIIFやTEIなどの国際標準規格に準拠して記述することで、第三者による多様な利活用を支援する形でデータ作成、研究資源の構築を行なった点である。

一方、構築コストの高さが課題として挙げられる。水滸伝は少なく見積もっても15,000行以上からなる作品であり（行数はバージョンによって大きく前後し、30,000行を超えるものもある）、上述したマークアップを人手で行うには限界がある。本研究では先行事例にならってテキストデータを行単位で作成しているため、例えば3.2.1については、2.2で挙げた既存のアプリケーションの成果やテキストマイニング技術等を活用し、既存のテキストデータ群から異同を半自動的に抽出するような手法が求められる。3.2.2については、画像処理技術等を活用し、画像からの行切り出しを自動化するような手法が求められる。

また、現状のOCRによるものでは精度に限界があるし、研究者が原資料を確認してテキストデータを作成することが基礎研究として重要であ

るため、現在、専門の研究者が手作業でテキストデータを作成している。その結果、テキストデータ自体の構築コストが莫大なものとなっている。今後の課題としたい。

## 4. まとめと意義

### 4.1 まとめ

2.2で紹介した通り、『水滸伝』をはじめとする白話小説のバージョン研究へのデジタル技術の活用には既に20年近い成果の蓄積があるが、従来はバージョン間の比較で使用するアプリケーションやプラットフォームが簡体字中国語版Windowsに制約されていた。IIIFやTEIを用いることでそれらの縛りから解放される可能性を開拓したのが本研究における第一の成果である。

テキストデータ自体も、先行事例では改行と改ページの位置情報を持つとは言っても基本的にはプレーンテキストだったものを、画像に紐づけたタグ付きテキストとして提供する道筋を示した。

挿絵の研究について、今回は方向性を示すにとどまったが、やはりタグ付が研究の進展の手助けとなるであろうとの見通しは得られた。

### 4.2 本研究の意義

本研究の第一の意義は、東京大学の図書館に所蔵されている資料に対して、図書館所属の研究部門であるU-PARLが基礎データ（JPEG画像、書誌データ、本文のテキストデータ）を作成し、公開する点にある。元来、図書館が提示する資料のデジタル画像は、図書館として求められる書誌データ（浅い記述）のみを付したものが多く、このような状況に研究者が関与することにより、研究者の必要とするレベルで書誌データを作成することができるばかりか、本文のテキストデータの付与も可能となるため、深い記述が可能となる。図書館員と研究者の協働という点で、U-PARLの試みは一つのモデルとなるだろう。

また第二の意義は、デジタル資料の作成を図書館という一組織で完結させるのではなく、東京大学デジタルアーカイブズ構築事業と協働で達成させることで、全学的なアーカイブ構築ないしは全学的な基盤の構築に寄与するとともに、IIIF画像ならびにTEIに準拠したタグ付けを行うことで、データが機械可読な形式で外部に開かれるという点にある。研究者が必要とする書誌データ（メタデータ）、本文のテキストデータ、IIIF準拠の高精細画像、TEI準拠のタグ付きテキスト、という基礎的な内容・技術の組み合わせではあるが、このような愚直なデータの公開こそ、息の長いデータとなるのであり、これこそ、図書館が取り組むに相応しいデジタル化作業だと言える。

## 参考文献

- [1]. 東京大学附属図書館：東京大学デジタルアーカイブズ構築事業，入手先  
<<https://www.lib.u-tokyo.ac.jp/ja/library/contents/archive-s-top>> (参照 2018-11-01).
- [2]. 中村覚，成田健太郎，永井正勝，富澤かな：U-PARL における漢籍・碑帖拓本デジタルアーカイブの試作と研究利用，情報処理学会研究報告，Vol. 2018-CH-116，No. 5，pp. 1-8 (2018) .
- [3]. 上原究一：商業出版に関わるテキストの問題——章回小説を例に，日本中国学会第2回次世代シンポジウム電子予稿，(2013) 入手先 <<http://nippon-chugoku-gakkai.org/utf8/jisedai/02/uehara.pdf>> (参照 2018-11-01).
- [4]. 周文業（著），小林瑞恵（訳）：『三国演義』 版本のデジタル化及び応用，三国志研究，第4号，pp.1-14 (2009) .
- [5]. Nicholas Laiacona, Ben Brumfield, Naoki Kokaze, Kiyonori Nagasaki, Makoto Goto : Connecting TEI and IIIF, IIIF Conference, (2018) .
- [6]. Versioning Machine, 入手先 <<http://v-machine.org/>> (参照 2018-11-01).
- [7]. 北本朝展，堀井洋，堀井美里，鈴木親彦，山本和明：時系列史料の人機分担構造化：古典籍『武鑑』を参照する江戸情報基盤の構築に向けて，人文科学とコンピュータシンポジウム じんもんこん 2017，pp. 273-280 (2017) .
- [8]. 小林宏光：中国版画史論，勉誠出版，東京，pp.107-232 (2017) .
- [9]. 上原究一：金陵唐氏世徳堂刊本講史小説三種と上元王氏の双面連式挿画について，瀧本弘之編『中国古典文学挿画集成（九）・小説集〔三〕』，遊子館，東京，pp.12-45 (2014) .
- [10]. 上原究一：孫悟空の図像イメージ——小説本文と絵姿と——，アジア遊学，171，pp. 82-102 (2014) .
- [11]. 上原究一：『李卓吾先生批評西遊記』の版本について，日本中国学会報，第63集，pp.110-126 (2011) .
- [12]. TEI Critical Apparatus Toolbox, 入手先 <<http://teicat.huma-num.fr/>> (参照 2018-11-01).
- [13]. TEIC/CETELcean: TEI in HTML5 Custom Elements, 入手先 <<https://github.com/TEIC/CETELcean>> (参照 2018-11-01).
- [14]. 人文学オープンデータ共同利用センター，IIIF Curation Viewer, 入手先 <<http://codh.rois.ac.jp/software/iiif-curation-viewer/>> (参照 2018-11-01).
- [15]. 鈴木 親彦，高岸 輝，北本 朝展：IIIF Curation Viewer が美術史にもたらす「細部」と「再現性」 絵入本・絵巻の作品比較を事例に，人文科学とコンピュータシンポジウム じんもんこん 2017，pp.157-164 (2017) .
- [16]. 人文学オープンデータ共同利用センター，IIIF Curation Finder, 入手先 <<http://codh.rois.ac.jp/software/iiif-curation-finder/>> (参照 2018-11-01).