

# 東方学 IIF 図書館： IIF 標準による漢籍デジタルアーカイブシステムの 構築

劉 冠偉

京都大学人文科学研究所

2025/07/25

## 概要

本研究は、京都大学人文科学研究所が長年にわたり公開してきた「東方学デジタル図書館」の近代化を目的とし、国際的な標準規格である IIF (International Image Interoperability Framework) に準拠した新たな漢籍デジタルアーカイブシステム「東方学 IIF 図書館」の設計と実装を報告するものである。現行システムは、IIF 未対応、検索機能の制限、モバイル非対応、永続的 URL の不在といった、現代のデジタル人文学研究の要求に応えられない複数の課題を抱えている。これらの課題を解決するため、本研究では Next.js と Nest.js を基盤とした 3 層アーキテクチャを採用し、Tiled Pyramid TIFF 形式の画像を IIPImage Server で配信するシステムを構築した。特に、既存の「全国漢籍データベース」とのメタデータ連携を図りつつ、IIF マニフェストを動的に生成する機能は本システムの核となる特徴である。本稿では、現行システムの課題分析から、IIF を適用した新システムの要求定義、アーキテクチャ設計、具体的な実装、そして今後の展望までを詳述する。本システムの構築は、漢籍資料の国際的な相互運用性を飛躍的に向上させ、デジタルアーカイブにおける標準化を推進するとともに、国内外の研究者にとってより利便性の高い研究基盤を提供することを目指すものである。

## 1 はじめに

デジタル人文学の発展に伴い、図書館、博物館などの機関が所蔵する貴重な資料のデジタル化と公開が世界的に進展している。特に漢籍をはじめとする東アジアの古典文献は、その学術的価値の高さから多くの研究機関でデジタル化プロジェクトが実施されている。京都大学人文科学研究所（以下、

人文研）もその先駆的な取り組みの一つとして、1997 年より「東方学デジタル図書館」を通じて漢籍画像データベースを公開してきた。

しかしながら、デジタルアーカイブの技術標準は急速に進化しており、相互運用性 (interoperability) の確保が重要な課題となっている。特に IIF (International Image Interoperability Framework) は、高品質なデ

デジタル画像の共有と活用を可能にするオープンスタンダードとして、国際的な図書館、博物館、アーカイブズ機関で広く採用されている。現行の「東方学デジタル図書館」システムは、IIIF への未対応をはじめとして、検索機能の制限、モバイル端末での利用困難、コマレベルでの URL 不在など、現代のデジタルアーカイブに求められる機能面で課題を抱えている。

本研究は、これらの課題を解決するため、IIIF 標準に準拠した新たな「東方学 IIIF 図書館」の開発を行うものである。本システムの構築により、漢籍資料の国際的な共有と活用の促進、研究者の利便性向上、そしてデジタルアーカイブにおける標準化の推進を目指す。

本論文では、まず既存の「東方学デジタル図書館」の現状と課題を整理し（第2章）、IIIF 標準の概要とその利点について述べる（第3章）。続いて、システム設計（第4章）と実装（第5章）について詳述し、最後に今後の展望を示す（第6章）。

## 2 東方学デジタル図書館の現状分析

### 2.1 システムの概要と歴史

「東方学デジタル図書館」<sup>1</sup>は、人文研が1997年より画像を公開している漢籍画像データベースである。2004年5月に現在の名称で正式公開され、プリンストン大学によって「Jinbunken Rare Books」としても

紹介されている(1)。現在、319点の漢籍と約11万4千枚の画像を公開しており、一部資料については高解像度画像も提供されている。

### 2.2 現行システムの構造

システムは、トップページから各漢籍の「冊・巻ページ」へ、さらに個別の「冊閲覧ページ」へと遷移する静的な階層構造で構築されている。閲覧ページでは、初期表示は低解像度画像であり、画像クリックによって高解像度画像に切り替わる。

東方学デジタル図書館の構造は以下の通りである：

```
top.html (トップページ)
├─ A029menu.html (冊・巻ページ)
│   └─ A0290001.html (冊閲覧ページ)
│       └─ 低解像度画像 (デフォルト)
│           └─ 高解像度画像 (クリック時表示)
```

この静的な HTML 構造は、利用者が段階的に資料へアクセスする流れを想定しているが、柔軟な検索や直接的なページ指定には対応できない。

### 2.3 課題の詳細分析

#### 2.3.1 技術的課題

**IIIF 標準への未対応:** 最大の課題は、国際標準である IIIF への未対応である。これにより、Mirador や Universal Viewer といった標準ビューアでの閲覧や、他機関の IIIF 対応リソースとの比較研究が不可能となっている。

**検索機能の制限:** 検索はトップレベルの書名に限定され、巻名や詳細な書誌情報

<sup>1</sup><http://kanji.zinbun.kyoto-u.ac.jp/db-machine/toho/html/top.html>

による絞り込みができない。多巻構成の漢籍において特定の巻を直接探すことができず、研究の効率を著しく妨げている。

**永続的 URL の不在:** 個別の画像 (コマ) を指し示す固定 URL が存在しないため、論文での引用や研究者間での情報共有が極めて困難である。これは学術研究における再現性と検証可能性を損なう重大な欠点である。

### 2.3.2 利用者体験の課題

**モバイル端末での利用困難:** レスポンシブデザインに対応しておらず、スマートフォンやタブレット端末での閲覧・操作性が著しく低い。研究環境が多様化する現代において、これは致命的な制約である。

**多言語対応の不足:** インターフェースが日本語のみであり、漢籍研究を国際的に行う海外の研究者にとって大きな障壁となっている。

**旧式のユーザーインターフェース:** 約 20 年前のウェブ標準に基づく UI は、現代の利用者にとって直感的でなく、学習コストが高い。これは特に若手研究者や学生の利用を妨げる要因となりうる。

### 2.3.3 相互運用性の課題

**全国漢籍データベースとの連携不足:** 同じ人文研が運営する「全国漢籍データベース<sup>2)</sup>」の豊富な書誌情報と、本システムの画像データが十分に連携していない。研究者は両システムを個別に検索・参照する必要があり、非効率を強いられている。

**権利情報の不明確性:** デジタル画像の利用条件やライセンスに関する情報が明確に

提示されておらず、研究者が画像を二次利用する際の判断を困難にしている。これはオープンサイエンス推進の観点からも問題である。

## 3 IIIIF 標準とその適用

### 3.1 IIIIF 標準の概要

IIIIF (International Image Interoperability Framework、トリプルアイエフ) は、高品質で属性付けされたデジタルオブジェクトを大規模にオンラインで配信するためのオープンスタンダードである(2)。

IIIIF のオープンな API 仕様は、異なる機関間での高品質なデジタル画像の共有と利用を可能にする。Image API(3)は画像の特定領域の切り出しやサイズ変更、回転などを URI で指定して取得できる。また、標準化されたマニフェスト (Presentation API(4)の JSON ドキュメント) により、異なるビューアやツールでのコンテンツの閲覧・利用が促進され、図書館、博物館、アーカイブズ機関が保有するデジタルコンテンツの共有と活用が進んでいる。

### 3.2 IIIIF 採用のメリット

IIIIF の採用により以下のメリットが期待される：

- 相互運用性: 異なるシステム間でのコンテンツ共有と再利用
- 高性能配信: マルチレゾリューション画像による効率的な配信
- 豊富なツール: Mirador、Tify 等のビューアの活用

<sup>2)</sup><http://kanji.zinbun.kyoto-u.ac.jp/kanseki>

- 標準化: 国際的なベストプラクティスへの準拠
- 将来性: 継続的な仕様更新と機能拡張への対応

### 3.3 国内外の実装事例

IIIF はすでに世界の主要な文化機関で採用されている。国外では Digital Bodleian (オックスフォード大学)<sup>3</sup>や Gallica (フランス国立図書館)<sup>4</sup>、国内では国立国会図書館デジタルコレクション<sup>5</sup>、e 国宝<sup>6</sup>、国文学研究資料館の国書データベース<sup>7</sup>などが代表的な事例として挙げられる。これらの先進事例は、IIIF が学術研究基盤として確固たる地位を築いていることを示している。

### 3.4 漢籍アーカイブへの適用検討

漢籍特有の要件として以下を考慮する必要がある：

1. 縦書きテキスト: 右から左への読み順への対応
2. 複数冊構成: 全集や叢書の階層構造の表現
3. メタデータ: 全国漢籍データベースとの互換性確保

<sup>3</sup><https://digital.bodleian.ox.ac.uk/>

<sup>4</sup><https://gallica.bnf.fr/>

<sup>5</sup><https://dl.ndl.go.jp/>

<sup>6</sup><https://emuseum.nich.go.jp/>

<sup>7</sup><https://kokusho.nijl.ac.jp/>

## 4 システム設計

### 4.1 要求定義

本システムの開発にあたり、第 2 章で明らかになった課題を踏まえ、以下の要求を定義した。

#### 4.1.1 機能要求

**IIIF 完全準拠:** Image API 3.0 および Presentation API 3.0 に準拠する。

**データ構造の近代化:** 「全国漢籍データベース」との互換性を持ち、コレクションから個別の葉 (コマ) に至る階層構造を明確にする。

**永続的 URL の提供:** 全てのコレクション、漢籍、冊、葉の各レベルで永続的な URL を提供する。

**動的マニフェスト生成:** データベースの更新に応じて IIIF マニフェストを自動的に生成・更新する。

**高度な検索機能:** 書誌情報に基づく詳細検索やファセット検索を可能にする Web API を実装する。

**管理機能:** 管理者が漢籍メタデータや公開状態を容易に管理できるバックエンドシステムを提供する。

#### 4.1.2 非機能要求

**ユーザビリティ:** モダンな UI と、モバイル端末に最適化されたレスポンスデザインを採用する。

**多言語対応:** 日本語、英語、中国語のインターフェースを提供する。

**パフォーマンス:** 11 万枚以上の画像を Tiled Pyramid TIFF と画像サーバーにより高速に配信する。

**運用・保守性:** Docker によるコンテナ化でデプロイメントを容易にし、GitHubでのソースコード公開により透明性を確保する。

## 4.2 システムアーキテクチャ

上記の要求を実現するため、下図に示す3層アーキテクチャを採用した。

**プレゼンテーション層 (Frontend):** Next.js を用いてサーバーサイドレンダリング (SSR) による高速なウェブサイトを構築。レスポンシブ UI、多言語対応、IIIFビューア (Tify) の統合を担当する。

**アプリケーション層 (Backend):** Nest.js を用いて RESTful API を構築。IIIF マニフェストの動的生成、検索機能、メタデータ管理など、システムのビジネスロジックを担う。

**データ層 (Data):** IIPImage Server が IIIF Image API に準拠した画像配信を担当。画像ファイルは Tiled Pyramid TIFF 形式でストレージに保存。メタデータは「全国漢籍データベース」と連携して管理する。

全体のトラフィックは nginx リバースプロキシによって各コンポーネントに振り分けられる。全てのコンポーネントは Docker コンテナとして独立して運用され、高いスケーラビリティと保守性を実現する。

### 4.2.1 データ構造設計

「全国漢籍データベース」との整合性を最優先とし、以下の5層からなる階層構造を設計した。

```
コレクション (文庫・分類)
├── 漢籍レコード
```

```
├── 漢籍レコード (子)
│   └── 冊 (物理的単位)
│       └── 葉 (個別ページ)
```

この構造により、人文研の漢籍分類体系を維持しつつ、IIIF が要求する詳細なメタデータ管理を両立させる。データの階層構造を「コレクション」を最上位とし、所属の文庫および人文研の漢籍分類に従う(5, 6)。IIIF マニフェストは、利用者が資料の全体像を把握しやすい「漢籍」単位と、物理的なまとまりである「冊」単位の二階層で提供する。これにより、柔軟な資料アクセスが可能となる。

## 5 システム実装

### 5.1 画像変換処理

既存の JPEG 画像を IIIF に適した Tiled Pyramid TIFF へ変換するため、vips ライブラリを用いた自動化スクリプトを開発した (GitHub<sup>8</sup>)。Intel Xeon Gold 5420+ を 2 基搭載したサーバー環境で処理を実行し、約 11 万枚の画像変換には I/O 性能の制約から約 2 週間を要する見込みである。

### 5.2 IIIF 画像サーバーの構築

IIPImage Server を Docker コンテナとして構築し (GitHub)、nginx リバースプロキシ経由で公開した。CORS 設定を適切に行うことで、外部サイトからの画像埋め込み利用にも対応している。これにより、以下

<sup>8</sup>[https://github.com/toyjack/tiled\\_tif\\_converter\\_docker](https://github.com/toyjack/tiled_tif_converter_docker)

の URL で IIIF Image API サービスが提供される。

<https://image.kanji.zinbun.kyoto-u.ac.jp/toho/iiif/>

### 5.3 マニフェスト生成システム

システムの中心機能である動的マニフェスト生成は、Nest.js 製 API によって実現される。まず、現行ウェブサイトの HTML 構造と「全国漢籍データベース」のデータを解析・抽出し、初期メタデータを構築した（変換スクリプト: GitHub<sup>9</sup>）。このメタデータに基づき、API はリクエストに応じて動的に IIIF Presentation API v3 準拠の JSON-LD マニフェストを生成・配信する。これにより、将来的な資料の追加や修正にも柔軟に対応できる。

### 5.4 検索・閲覧インターフェース

Next.js を用いて開発した検索・閲覧インターフェースは、第 4 章の要求定義に基づき、書名・著者による全文検索、ファセット検索、レスポンシブデザイン、そしてカスタマイズした Tify ビューアによる快適な閲覧体験を提供する（図 1、2）。開発プロセスにおいては、GitHub Copilot 等の AI 支援ツールを活用し、特に UI コンポーネントの実装を効率化した。



図 1: 開発中の UI - ホームページ

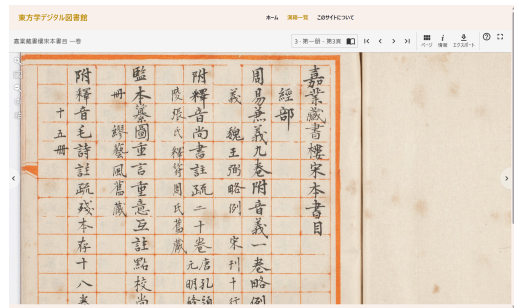


図 2: 開発中の UI - 閲覧ページ

## 6 今後の展望

本研究では、IIIF 標準に準拠した「東方学 IIIF 図書館」の設計と実装を行った。システムは現在部分的に稼働しており、画像変換作業も進行中である。今後の展望として以下の項目が挙げられる。

### 6.1 OCR と TEI 導入による全文検索

将来的な展望として、OCR（光学文字認識）の導入による全文テキスト検索機能の実装が期待される。漢籍の古典中国語テキストに対応した OCR エンジンの活用により、画像内の文字情報をテキスト化し、より精密な検索が可能となる。

<sup>9</sup><https://github.com/toyjack/toho-html-data>

## 6.2 アノテーション機能の追加

IIIF 標準のアノテーション機能を活用し、以下の機能を実装する予定である：

- 研究者による注釈の追加・共有機能
- 翻刻テキストの重ね合わせ表示

これにより、単なる画像閲覧システムを超えた、協働的な研究プラットフォームとしての発展が期待される。

## 6.3 システムの評価と改善

本システムの本格運用開始後は、以下の観点から評価を行い、継続的な改善を図る予定である：

- IIIF 準拠度の検証と国際標準への追従
- 利用者からのフィードバックに基づくインターフェース改善
- アクセスログ分析による利用実態の把握と最適化
- パフォーマンス指標の継続的な監視と改善

## 6.4 国際的な連携の拡大

IIIF 標準の採用により、以下のような国際的な連携の拡大を目指す：

- 海外機関との漢籍資料の相互利用促進
- IIIF Consortium との連携強化
- 多言語対応の充実（英語、中国語、韓国語等）

## 7 おわりに

本研究では、京都大学人文科学研究所の「東方学デジタル図書館」を IIIF 標準に準拠

したシステムに刷新する「東方学 IIIF 図書館」の開発について報告した。

現行システムの課題分析から始まり、IIIF 標準の特性を活かしたシステム設計、そして具体的な実装まで、包括的なアプローチを採用した。特に、全国漢籍データベースとの連携、動的のマニフェスト生成、レスポンスなユーザーインターフェースの実現により、従来システムの限界を克服することができた。

IIIF 標準の採用により、漢籍資料の国際的な相互運用性が大幅に向上し、研究者にとってより利便性の高いデジタルアーカイブシステムを構築することができた。また、オープンソースによる実装と Docker によるコンテナ化により、他機関での導入や応用も容易になった。

今後は、OCR と TEI 導入による全文検索機能、アノテーション機能の充実、国際的な連携の拡大などにより、単なるデジタルアーカイブを超えた研究プラットフォームとしての発展を目指す。本研究の成果が、デジタル人文学分野における IIIF 標準の普及と、東アジア古典文献研究の国際的な発展に寄与することを期待する。

## 参考文献

- (1) Noguchi, Setsuko. “East Asian Studies Links: Japanese”. <https://libguides.princeton.edu/eallinks/japanese>, (参照 2025-07-07).
- (2) IIIF Consortium. “International Image Interoperability Framework”. <https://iiif.io/>, (参照 2025-07-07).

- (3) Michael Appleby, Tom Crane, Robert Sanderson, Jon Stroop, Simeon Warner. IIF Image API 3.0. 2020. <https://iif.io/api/image/3.0>, (参照 2025-07-07).
- (4) Michael Appleby, Tom Crane, Robert Sanderson, Jon Stroop, Simeon Warner. IIF Presentation API 3.0. 2020. <https://iif.io/api/presentation/3.0>, (参照 2025-07-07).
- (5) 井波陵一. 京都大學人文科學研究所漢籍分類一覽：部-類-屬-目-例. 京都大學人文科學研究所付屬漢字情報研究センター, 2005. <http://hdl.handle.net/2433/65024>.
- (6) 安岡孝一. “全国漢籍データベースの設計とWWWでの運用”. <http://kanji.zinbun.kyoto-u.ac.jp/~yasuoka/publications/2002-11-19.pdf>, (参照 2025-07-07).