

A decorative background featuring a repeating pattern of stylized, rounded leaves in a lighter shade of purple against a darker purple background. The leaves are arranged in a way that creates a sense of movement and depth.

FROM CONCEPT TO REALITY

The establishment of ToUDA at
Tohoku University

Tomoe HANZAWA

From concept to reality

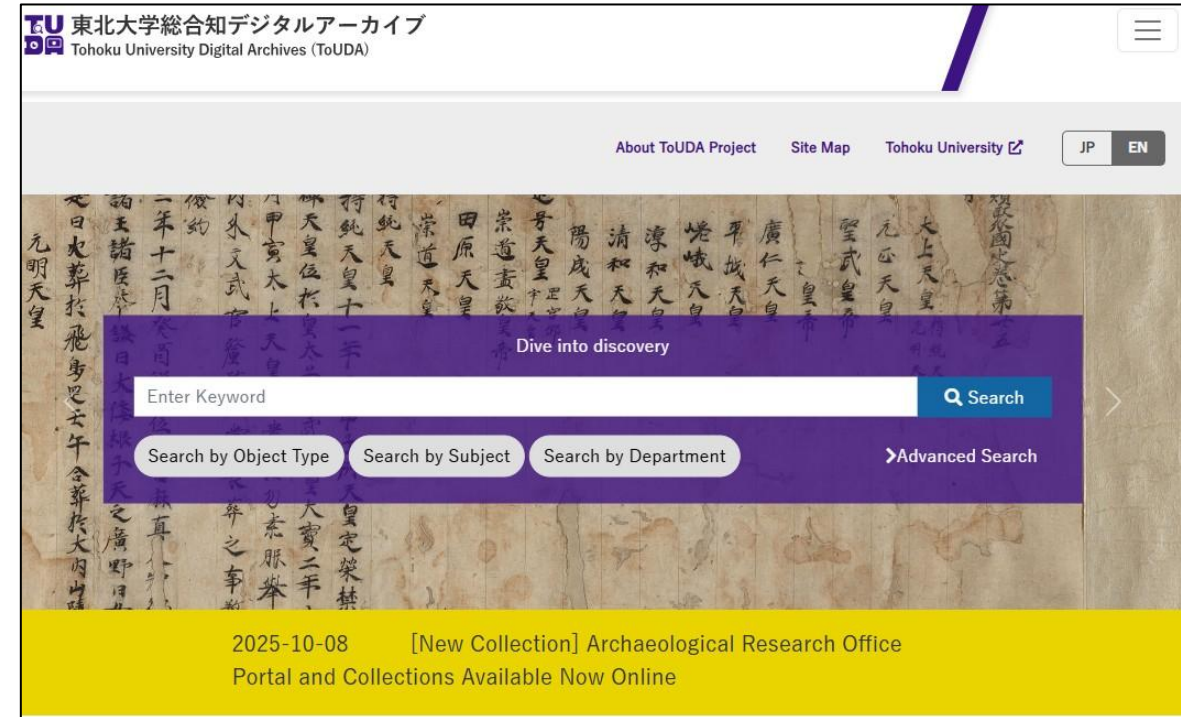
The establishment of ToUDA

Welcome and Thank You for Joining

Today, I will share the journey of how Tohoku University's digital archive project — ToUDA — evolved from a conceptual idea into a functioning system, and where it is headed next.

Purpose of This Talk

- To introduce the vision behind ToUDA
- To share its current status and achievements
- To reflect on the challenges encountered
- To explore future possibilities



- **Extensive Career in Library Services**

Over 31 years of experience working in various library roles including journals contracts and user guidance.

- **Academic and Professional Background**

Majored in Japanese history (specializing in early modern Japan) during university and graduate school

- **Leadership and Digital Services**

Currently directing user services, she manages digital archives and exhibitions, enhancing library accessibility.

Regarding Tohoku University and the Tohoku University Library, please refer to the <https://youtu.be/SOIkfIB9rbs>.

There is an introduction, so please take a look there.





VISION

What is ToUDA?

5

ToUDA = Tohoku University Digital Archives

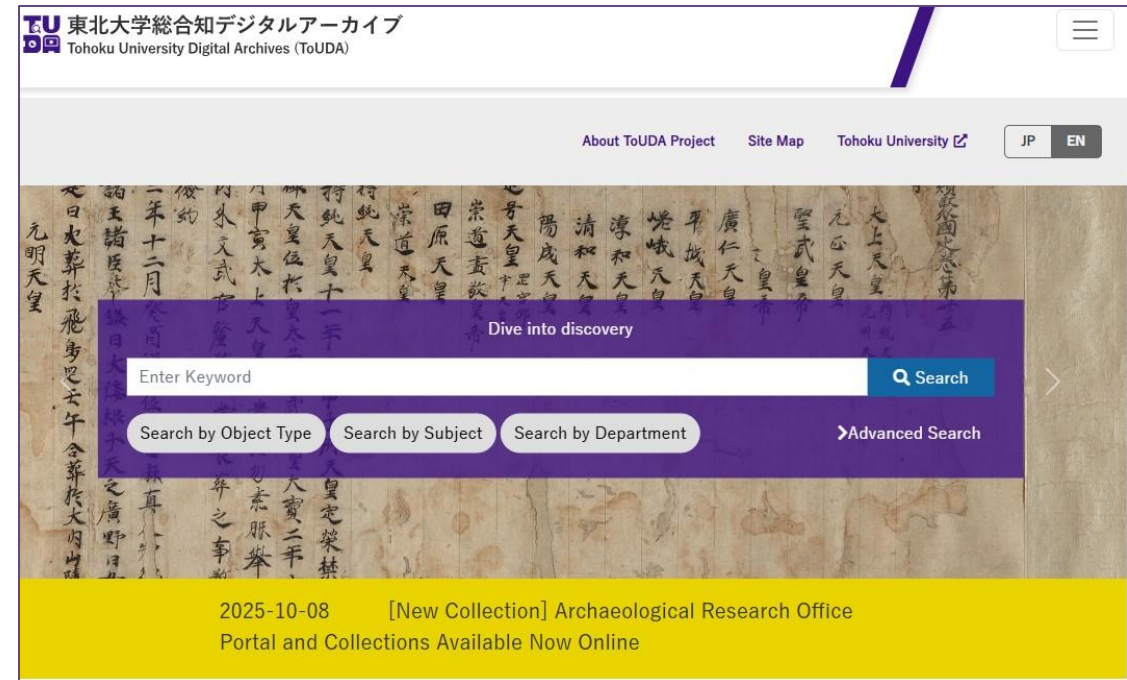
- Officially launched: April 4, 2024
- Japanese name: 東北大学知の統合デジタルアーカイブ
- URL: <https://touda.tohoku.ac.jp/portal/en>

What it includes

- Digitized materials from libraries, archives, and research institutes
Classical books, historical manuscripts, photographs, and more

Purpose

- University-wide platform for publishing and integrating academic and cultural resources, supporting interdisciplinary research and global access



<https://touda.Tohoku.ac.jp/portal/en>

Challenges We Faced

- Fragmented access to academic and cultural resources across departments
- Lack of a unified platform for interdisciplinary research
- Limited visibility of valuable materials beyond Japan
- Difficulty in maintaining and sharing metadata across systems

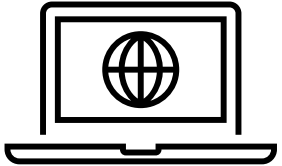
Why a University-Wide Digital Archive Was Needed

- To integrate and globally publish diverse resources
- To support cross-disciplinary knowledge creation
- To provide a shared infrastructure for departments without their own databases
- To enhance accessibility for researchers worldwide

Vision Behind ToUDA

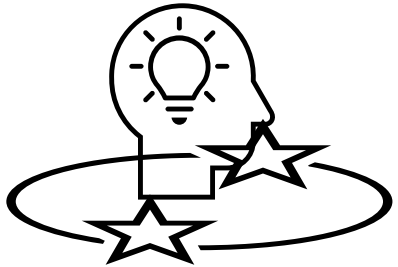
A comprehensive, interoperable, and globally accessible platform for academic and cultural heritage.

1. A Comprehensive Platform for Global Access



- Digitally publishes academic and cultural resources from across the university
- Complies with IIF standards for interoperability
- Offers bilingual access (Japanese & English)

2. Foundation for Integrated Knowledge Creation



- Designed to support interdisciplinary research
- Enables cross-sectional and large-scale analysis

3. Inclusive Infrastructure



- Provides digital archive environments for departments without their own systems
- Aggregates metadata from existing departmental archives
- Builds a university-wide portal for unified access

Metadata Design

- Based on paper materials
- Aligned with Dublin Core and National Diet Library standards
- Designed to be simple yet comprehensive

Search Functionality

- Departmental DBs: refined search tailored to each collection
- University-wide portal: filtering by field and material type, Full-text search available

Additional Features

- Full-text content provision
- Downloadable images of entire materials



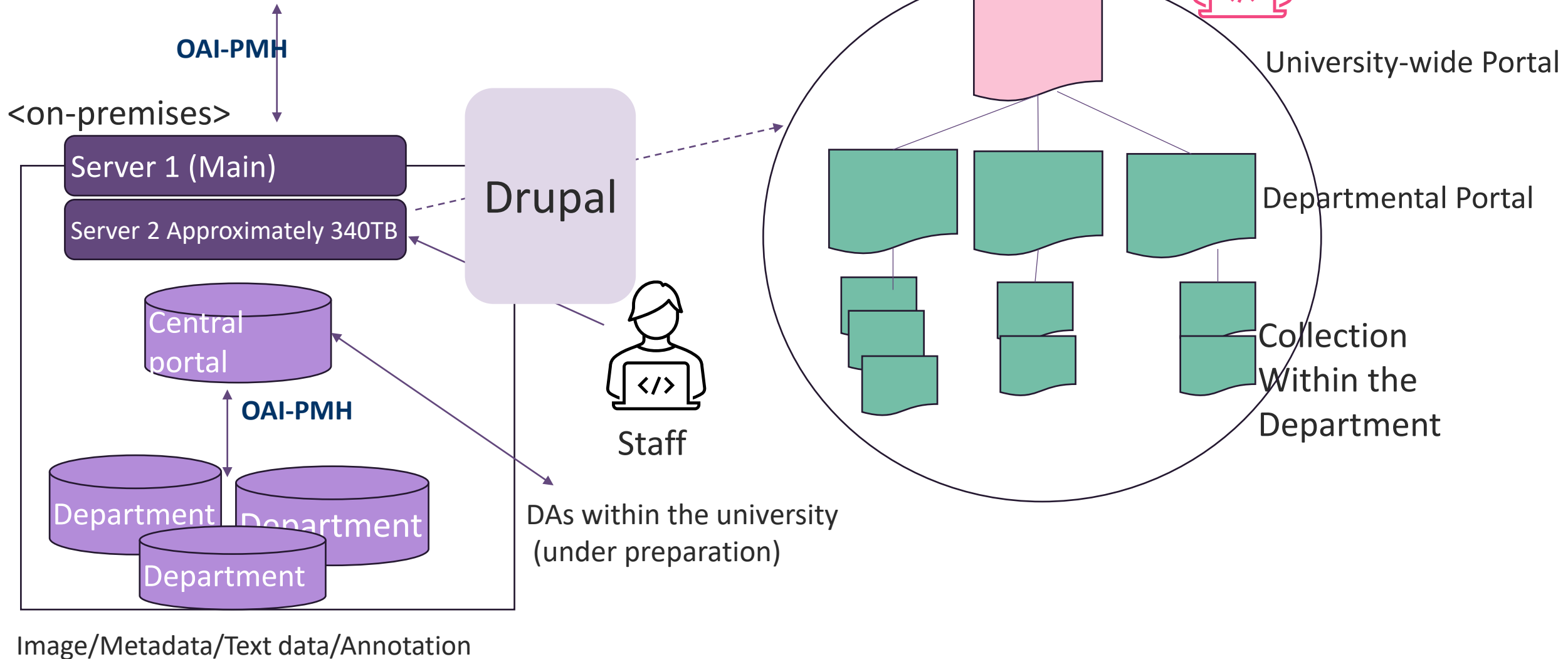
CURRENT STATUS

System Overview of ToUDA

10

Linkage

- **PRL**
- JapanSearch NDL Search (Under Preparation)



1 University Archives



※The numbers are approximate

with image		metadeta
13,000	/	90,000

2 Main Library



with image		metadeta
47,600	/	67,000

3 Medical brary



with image		metadeta
96	/	96

4 Engineering Library



with image		metadeta
1	/	1

5 Archaeological Research office on the Campus



with image		metadeta
181	/	181

In preparation

- Graduate School of Arts and Letters
- Graduate School of Law

- Number of Items with Images ※The numbers are approximate. *Italic indicates up coming.*

<div>Departments</div> <div>Material Types</div>	University Archives	Main Library	Medical Library	Engineering Library	Archaeological Research office on the Campus	Departments in preparation
Classical Books & historical documents-Japanese, Chinese		27,900 <i>400</i>	96			<i>400</i>
Classical books-Western				1		
Drawings, maps, paintings, photographs	7,300	17,300				<i>200</i>
Excavated artifacts					2D :181 <i>3D : the numbers has not been defined</i>	
Modern publication		1,000				
Modern and Contemporary Manuscripts & Documents (University Historical Materials, etc.)	5,700	1,400				
Sutra (Tibetan Tripitaka)		<i>200</i>				

Data in Preparation

(Scheduled for Release by the End of March Next Year)

- **Library**
 - Japanese historical documents (∼ 440 items)
 - Sutras (∼ 200 items)
- **Graduate School of Literature**
 - Japanese historical documents and glass plate photographs (∼ 800 items)
- **Graduate School of Law**
 - Japanese historical documents (13 items)
- **Archaeological Research Office on the Campus**
 - 3D data (number of items undecided)

Classical Books & historical documents-
Japanese, Chinese

Classical books-Western

Drawings, maps, paintings, photographs

Excavated artifacts

Modern publication

Modern and Contemporary Manuscripts &
Documents (University Historical Materials,
etc.)


Sutra (Tibetan Tripitaka)

Common Schema

- Designed to balance comprehensiveness and simplicity
- Compatible with Dublin Core and National Diet Library standards
- Number of fields: 138

Special Schema for Sutras

- Unique bibliographic structure required
- Two types of schemas:
 - Scripture texts and images
- Number of fields : for texts- 156, for images- 132



CONCEPTS vs. REALITY: IMPLEMENTATION STATUS

COMPONENT	STATUS
Global publication platform	Completed
DA environments for departments	Completed
University-wide portal	Completed
Metadata collection from departmental DAs	In Progress
Collaboration with external archives	In Progress
Metadata translation and descriptions	Preparation

Original Concept

- One common schema for paper based materials
- Reasonably comprehensive but not overly detailed
- Aligned with Dublin Core and National Diet Library Standard

Challenges Encountered

- Difficulty representing hierarchies and parallel relationships
- Operational complexity due to diverse cataloging methods
- Custom fields needed for research-specific materials (e.g., scriptures)
- Inconvenience in applying common display/hide settings and facets



Metadata schema in use, with ongoing adjustments

Original Concept

- Departmental DBs: refined search tailored to each collection
- University-wide portal: filtering by field and document type
- Full-text search across the metadata for deeper access

Challenges Encountered

- Difficulty setting facets for each collection due to cataloging differences
- Broad categories in the portal are not sufficiently useful for precise filtering
- Full-text search errors due to bulk migration (e.g., kanji misreadings)



- Basic search functions are operational
- Refinement and customization are ongoing
- Need for improved metadata structure and search logic

ORIGINAL CONCEPTS	CURRENT STATUS	CHALLENGES
Provide Full-text	Implemented	High accuracy requires time and cost
Bulk image download	Provided via compressed files upon request	Server capacity needs to be secured



- Planning to input within the range of automatic generation
- Expansion planned in the next system update



FROM NOW ON

CHALLENGE	CURRENT STATUS	FUTURE PLAN
Scale	-Museum materials not yet included; only 2D images and minimal metadata	-Consider methods to record detailed data and build systems for access and use
Search Function	-Conventional search tools are insufficient for diverse materials	-Redesign search functionality with AI to improve efficiency and usability
Data diversity	-Basically single metadata scheme -Little data other than bibliographic data	-Designing schemas for each material type -Richen the data

Looking Ahead

— Toward a Research-Driven Archive : toward a dynamic research infrastructure

Strengthening Researcher Engagement

We will enhance communication with researchers to better understand their needs and reflect them in system improvements.

AI-Driven System Enhancement

AI technologies will be applied to improve search precision, metadata quality, and overall system efficiency.

Update Timeline

Detailed design work will begin in April 2026, with a major system update scheduled for April 2028.

The background is a solid dark purple. In the corners, there are stylized, lighter purple leaf and branch patterns. The top-left and bottom-left corners feature clusters of small, rounded leaves on thin stems. The top-right and bottom-right corners feature larger, more detailed leaves with visible veins and stems, some of which are curved. The text "Thank you" is centered in the middle of the image.

Thank you