

# 中尾佐助スライドデータベースと大学リポジトリとの連携に向けて Toward System Cooperation between Sasuke Nakao Slide Database and Institutional Repository

小島 篤博, 青木 茂樹, 泉 正夫, 宮本 貴朗

Atsuhiko Kojima, Shigeki Aoki, Masao Izumi, Takao Miyamoto

大阪府立大学大学院 人間社会システム科学研究科, 大阪府堺市中区学園町 1-1

Osaka Prefecture University, 1-1, Gakuen-cho, Nakaku, Sakai, Osaka

**概要:** 大阪府立大学では、故中尾佐助博士がフィールド調査において撮影した資料をデータベース化した中尾佐助スライドデータベースを開発、公開している。2000年の公開以来、数度のシステム更新を経て現在に至るが、基本的な設計は開発当初から大きく変更されておらず、長期的な運用が懸念されつつある。一方、近年ではデジタル資料の機関リポジトリへの集約により、永続的なアーカイブ化が進められている。本報告では、中尾佐助スライドデータベースのこれまでの運用状況とともに、大学リポジトリとの連携を考慮した今後の展望について紹介する。

**Abstract:** Library and science information center of Osaka Prefecture University maintains a slide image database taken by the late Dr. Sasuke Nakao. Since 2000 when the system was made to the public, it has been updated several times as the basic structure remains unchanged. In recent years, such research materials are being gathered on persistent digital archives in the institutional repositories. This paper describes the present state of the Sasuke Nakao slide database as well as a prospect of the system cooperation between the database and the University's institutional repository.

**キーワード:** 中尾佐助, 照葉樹林文化論, 画像データベース, 機関リポジトリ

**Keywords:** Sasuke Nakao, theory of evergreen forest culture, image database, institutional repository

## 1. はじめに

大阪府立大学学術情報センターには、照葉樹林文化論[1]の提唱者として名高い故・中尾佐助大阪府立大学名誉教授の旧蔵書、雑誌、スライド、遠征アルバム、フィールドノート、スクラップブックなど膨大な研究資料(中尾佐助コレクション)が保存され、大阪府立大学の特色あるコレクションの一つとなっている。これらのうち写真やスライド、フィールドノートなどについては、スキャナにより画像データとして読み取り、Web上で閲覧可能な中尾佐助スライドデータベースとして公開されている。

現在公開の対象となっているのは、中尾資料のうち22,000枚のスライドであり、これに撮影場所や日時などのメタデータを付与した上で、キーワードや場所、日時などにより検索することができるようになっている。学術的・文化的な研究資料をデジタルアーカイブ化し、

インターネット上に公開する試みは他にもなされているが、本システムの特徴は、研究者が個々の画像ごとに独自のキーワードやコメントを付与する機能を有することである。大阪府立大学では1997年よりこのデータベースの開発を進め、公開から約20年を経た現在までに幾度かシステム更新を行ってきたが、基本的な設計は当初のものをそのまま踏襲している。研究機関が独自に開発したシステムは、研究者個人の努力によるものが多く、異動や退職により運用の継続が難しくなるなどの懸念を抱えており、人文系データベースの分野においてもメンテナンスを中断せざるを得なかった例も聞かれる。

一方、近年では、大学などの研究機関が所蔵する論文や図書、研究資料等を機関リポジトリとして集約する動きが進んでいる。中尾佐助資料に関しても、資料の一部であるフィールドノートが大阪府立大学学術情報



図1 中尾佐助スライドデータベース

リポジトリ OPERA に収容されている。機関リポジトリに収容された資料は、永続的な URI を与えられ将来に渡って利用可能であることが期待できる。したがって、前述の独自開発型データベースと機関リポジトリを横断的に連携することで、長期的にもデジタルアーカイブとして利用できることを担保しつつ、その時代に応じた独自機能の追加を実現することができると考えられる。

本論文では、中尾佐助スライドデータベースのシステム構成について紹介するとともに、大学リポジトリとの連携により、デジタルアーカイブとしての継続性と独自機能の実現の両立を目指した将来構想について述べる。

以下、2では中尾佐助資料の概要について紹介し、続いて3では現行システムの概要について説明する。4では、データベースの構成とシステムの具体的な機能について説明し、5で大学リポジトリの連携について展望する。最後に6でまとめる。

## 2. 中尾佐助資料について

大阪府立大学が保管する中尾佐助コレクションは、旧蔵書 3,500 冊、雑誌 60 タイトル、著書 18 点、学術論文 88 点、新聞・雑誌等の記事 552 点、スライド約 28,000 点、遠征アルバム 77 冊及び写真ネガ、スクラップブック 94 冊、記録ノート 116 冊、フィールドノート 19 冊等からなり、中尾佐助コーナーにおいて閲覧に供している[2]。このうち、スライドフィルムは 1955 年から 1984 年の海外調査において撮影されたものである。遠征アルバムには、1935 年頃から 1950 年代前半のカラーズライドにするまでものが白黒写真で残されている。

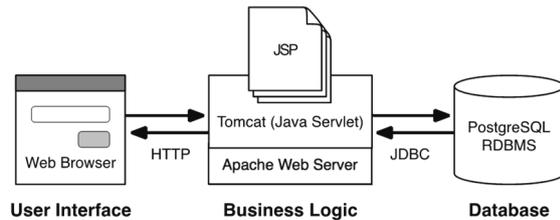


図2 システム構成図

写真帳の写真の番号とネガの番号が併記しており、個々のスライドを特定できるようになっている。

大阪府立大学学術情報センターでは、これら中尾資料のデジタルアーカイブ化を進めている。まず、1999 年より第 1 期開発として、中尾が 1958 年にブータンのフィールド調査を行った際のスライド画像 1,300 枚を対象とし、個々のスライドにキーワードや場所、日時などを関連付けることで、キーワードのほか地理的、時間的條件から目的の画像を検索することができる画像データベースを構築し、公開してきた[3][4]。これに続く第 2 期開発として、2006 年度より、残りのカラーズライド画像についてもキーワード付与などの作業を行い、第 1 期公開分と合わせて約 22,000 枚の画像データベースとして 2009 年より公開している(図 1)。後述する研究者支援機能は、この第 2 期において実装したものである[5]。

さて、これらのスライドフィルムの形態としては、個々のスライドのマウントに、1 枚ごとに識別番号が付与されている。この識別番号をもとに、フィルムごとに分類整理され、保管されている。例えば、“N58-54-12”は、1958 年に撮影された、54 番目のフィルムの 12 コマ目となる。この他、スライドの情報を得る手がかりとしては、中尾自身の著作である『秘境ブータン』[6]、『ヒマラヤの花』[7]、『ブータンの花』[8]があり、フィールドノート(3 冊)とブータンのアルバム(12 冊)が揃っていた。また、フィールドノートには、キャンプ地と標高、キャンプ地からキャンプ地への移動行程や天候等が明記されており、スライドを特定する際に有用な情報を得ることができた。

## 3. 現行システムの構成

### 3.1 システムの設計

現行の中尾佐助スライドデータベースは、以下の機能を有している。

- (1) ユーザ登録・管理
- (2) さまざまな条件によるスライドの検索
- (3) コメントやキーワードの付与
- (4) ブックマークの付与

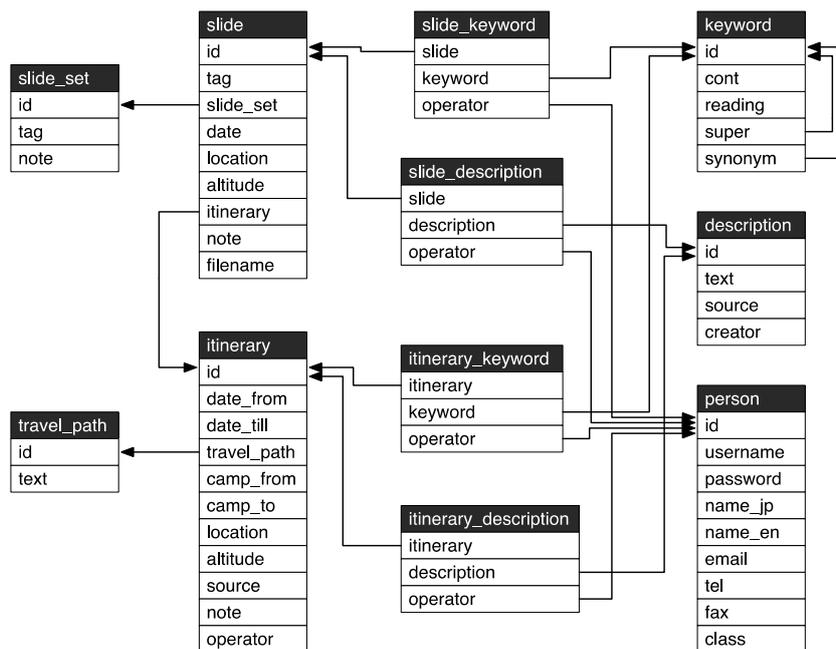


図3 データベースのE-R図

表1 スキーマ一覧

スキーマ	内容
slide	スライドの基本情報
slide_set	スライドセット単位の情報
itinerary	フィールド調査の旅程
keyword	キーワード情報
description	文献のテキスト
slide_keyword	スライド=キーワード対応表
slide_description	スライド=テキスト対応表
person	登録ユーザ情報
comment	ユーザが付与したコメント
bookmark	ユーザが付与したブックマーク

まず、ユーザごとに作業履歴等を保存するため、研究者向け機能の利用にあたってはユーザ登録をいただいている。ユーザは、利用開始時にパスワードによる認証を行い、以後の作業はそのユーザに関連付けて記録される。なお、研究者専用機能を利用しない場合は、ユーザ認証なしで検索や画像表示などの基本的な閲覧機能が提供されている。

スライドの検索は、キーワード、日付、番号、および一覧からの選択のいずれかの方法により行うことができる。スライドを特定した後、必要があれば、そのスライドにコメントやブックマークを付与することができる。コメントは、スライドの内容について自由記述形式で記入ことができ、公開の可否も選択可能である。また、ブックマークに登録しておけば、あとで簡単にアクセスすることができる。

以上のような要件を考慮してシステム構成を検討した結果、図2に示すような3階層型システム(3-tier system)として実現することとした。これにより、一般ユーザからのスライド検索や、特定ユーザによるデータ編集作業などはWebインタフェースにより行い、そのためのビジネスロジックはWebサーバ上に実装し、さらに基本データ構造やそれらの関連性はRDBMSに格納することができる。

Webサーバの構成はApacheとTomcatにより実装し、動的なページ生成にはJSP(JavaServer Pages)を利用している。JSPでは、サーバ側の処理をJava言語で記述できるため、Webからの要求に応じて動的にSQLを組み立てて、RDBMSにクエリを投入するなど、きめ細かな処理を行うことが可能である。バックエンドのRDBMSとしてはPostgreSQLを採用しており、JDBC(Java Database Connectivity)を介してTomcat上のJavaと接続している。また、ODBC(Open Database Connectivity)を介してMicrosoft Accessなどのデータベース・フロントエンドと接続することで、容易にデータの編集が可能となっている。

### 3.2 データベースの定義

スライド画像をキーワードから検索するには、適切なキーワードをスライドに対応付ける必要がある。また、キーワードやコメントを付与した作業者を記録するためには、ユーザ情報についてもデータベースに格納する必要がある。本研究では、これらの項目を組み合わせ、



図 4 キーワード検索画面



図 5 スライド編集画面

1 つの検索システムとして機能するようデータベースのスキーマ設計を行った。本システムのスキーマの関連を表した E-R 図を図 3 に、また、スキーマの一覧を表 1 に示す。

ここで、slide は個々のスライドの基本情報(撮影日時、場所、画像ファイル名など)、keyword はキーワード情報(キーワード、英表記、同義語へのリンクなど)、また person は登録ユーザ情報(ユーザ名、パスワード、氏名、連絡先など)を保持している。

### 3.3 スライドの検索

キーワードによるスライドの検索は、ユーザが入力したキーワードと適合するレコードを keyword テーブルから検索し、そこから slide\_keyword を辿ることで目的の slide を選択する。また、複数のキーワードを指定する場合は、それぞれ選択されたスライドの集合に対し、積集合や和集合をとることができる。

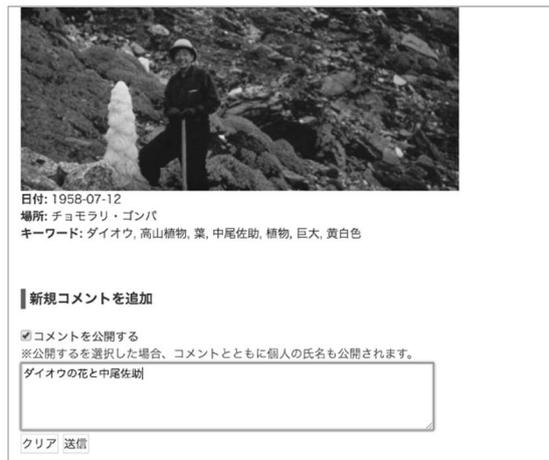


図 6 コメント編集画面

実際のキーワード検索画面を図 4 に示す。キーワードとして利用できる文字は、漢字、ひらがな、全角カタカナ、英数字である。また、個々のキーワードに対して、前方一致、後方一致、部分一致、完全一致、アルファベットの大小文字の区別も指定できるようになっている。キーワードを指定して「検索」をクリックすると、該当するスライドのサムネイル画像が一覧表示され、さらにサムネイル画像を選んでクリックすると、スライド画像とともにキーワード、コメントなどの付随情報が表示される(図 5)。

### 3.4 データの編集

スライドの基本情報として、キーワードやコメントを追加することができる。キーワードを付与するためには、まず対象のスライドを選択する。例えば、先述の検索により選択されたスライド画面の下には編集メニューが表示され、ブックマーク、キーワード、コメントを付与できるようになっている。ここで「キーワードを付与」というリンクをクリックすると、キーワード編集画面が表示され、新たなキーワードの追加や既存のキーワードの削除ができる。なお、既存のキーワードは、付与者自身または管理者によってしか削除できないように制限されている。

新たに入力したキーワードが、既存の keyword のテーブルに存在しなければ、新たに keyword のレコードを追加する。同時に、slide\_keyword のテーブルにレコードを追加し、それぞれ slide と keyword との関連付けを行う。

コメントの付与についても、キーワード同様、対象スライドを選択したのち編集メニューから行うことができる。「コメントを付与」をクリックすると、図 6 のようなコメント編集画面が表示される。この画面は、コメントを公開す



索引					
頁数	摘要	頁数	摘要	頁数	摘要
6	0900-1000	26	11,990	25	13
7	1800-1900	70	18,235	19	47
6	1300-1500	70	11,705	25	6
7	2200-2400	70	18,235	19	8
6	1000-1200	70	11,705	25	6
		70	18,235	19	8
6	1500-1600	70	18,235	25	9
7	0000-0100	70	18,235	21	5

図7 大阪府立大学リポジトリ

るかどうかのチェックボックスと、自由記述形式でコメントを入力できるテキストボックスから構成されている。それぞれ入力し、「送信」ボタンをクリックすれば、対象スライドにコメントを付与することができる。

## 4. 大学リポジトリとの連携

### 4.1 現行システムの課題

ここまで現状のシステム構成について述べてきたが、将来的な課題として、いかに本システムのサービスを長期的に継続していくかという懸念がある。研究資料のデジタルアーカイブは永続的なアクセスが期待される一方、独自開発されたシステムでは OS やサーバソフトウェアのサポート期限や、ミドルウェアの開発停止など寿命がある。多大な労力を費やして構築したデータベースも、一定期間ごとにシステムを更新し続けなければ、やがて利用できなくなってしまう。

本システムも、当初は Linux 2.2 上に Java ベースの Web サーバである Jigsaw 2 を基盤に Java Servlet として実現し、DB サーバは Windows NT 4.0 上の Oracle 8 に接続する 2 サーバ構成で構築した。しかしながら、Jigsaw の開発が停止し Java ベースの Web アプリケーション基盤は Apache Tomcat プロジェクトに移行したため、第 2 期では RedHat Enterprise Linux 5、Apache 2.0、Tomcat 6、そして RDBMS は PostgreSQL 8 に更新した。現行システムでは CentOS 7、Apache 2.4、Tomcat 7、PostgreSQL 9 に更新している。Web アプリケーションのトレンドは現在も変化しており、PHP のほか Ruby

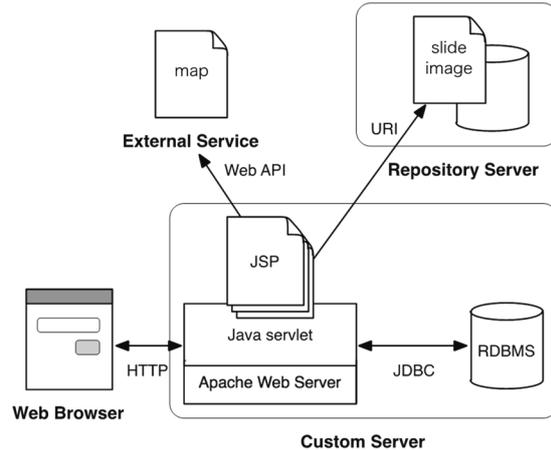


図8 リポジトリとの連携

on Rails や Python が普及し、更には開発当初には存在しなかったマッシュアップ型の様々なサービスが登場してきている。

蓄積したデータを将来に亘って安定的に提供し続けるためには、より汎用的・公共的な基盤を利用するのが確実であると考えられる。近年、大学や国立研究機関で集約が進められている機関リポジトリは、まさに研究資料を蓄積する目的で開発されたものである。大阪府立大学においても、オープンソースのデジタルアーカイブシステムである DSpace[9]を採用した学術情報リポジトリ OPERA を運営している[10]。このリポジトリには中尾コレクションの一部も収容されており、図 7 にフィールドノート(1958年8月16日-10月26日):ブタンの掲載ページを示す。DSpace は多くのデータ形式をサポートしており、画像データをそのメタデータとともに蓄積・管理できるため、本システムのスライド画像を移行することができれば、長期的に安定したデータの管理とサービスを維持することが期待できる。一方、検索方式は画一的であり、基本的にはテキストのパターンマッチングのみとなってしまうため、現在本システムが提供しているような地図から撮影場所を絞り込んだり、撮影日時で検索したりするなどの独自の機能は実現できない。

### 4.2 将来の構想

そこで、本システムの将来構想として、画像データやそれに付随する基本情報は大学リポジトリに収容し、独自の検索機能は別システムとして分離して、2 サーバを連携することでシステムを実現することを検討している。図 8 にそのシステム構成を示す。現行のシステムに蓄積されているスライド画像データとその基本情報は、リポジトリに収容可能なデータ形式に変換して移行する。その際、個々のスライドに付与されたキーワ

ードは、テキスト形式で出力してメタデータの一部とする。リポジトリがサポートしない独自の検索、例えば地図や日時からのスライド検索はカスタム検索サーバ上でを行い、検索結果のスライド画像のみリポジトリ上のURIにリンクする形でページを生成する。

これにより、スライド画像を含む基本的なデータは大学リポジトリ上で長期的な参照を担保しつつ、独自機能は別サーバに分離し、万が一サービスを継続できなくなった場合でも基本的なデータだけはリポジトリ上に残すことができる。

更には、カスタム検索サーバを実現する際、地図上の位置の参照は Google Maps API など外部サービスを利用したり、キーワードは Linked Data[11]や RDF に準拠することで汎用性を高めたりするなど、独自開発が必要な部分を徐々に縮小していくことも考えられる。

以上の構想に基づいて、現在はシステムの基本設計を行なっている段階であり、基本機能の分離を実現したパイロットサーバを開発する予定である。

## 5. おわりに

大阪府立大学が保管する中尾佐助資料を Web 上で検索可能な中尾佐助スライドデータベースと、その継続的な利用を目指した大学リポジトリとの連携について述べた。本システムは以下の URL で公開しており、研究者登録をしていただければ、キーワードやコメントを付与することもできる。関連分野の研究者にご利用いただければ幸いです。

<http://nakao-db.center.osakafu-u.ac.jp/>

## 謝辞

本研究において、データベースの要件および機能改善について有益なご助言いただいた山口裕文大阪府立大学名誉教授、ならびにデータベースの運営に日々ご協力いただいている大阪府立大学学術情報室の皆様にご感謝の意を表します。

## 参考文献

- [1] 中尾佐助. 中尾佐助著作集 第IV巻 照葉樹林文化論. 札幌, 北海道大学出版会, 2006.
- [2] 大阪府立大学総合情報センター編. 中尾佐助文献・資料総目: 照葉樹林文化論の源流. 堺, 大阪府立大学総合情報センター, 1997.
- [3] 富坂敏子, 小池利栄子, 小島篤博, 石井敬三, 宮本貴朗, 山野美贇子. 中尾佐助資料スライドデータベースの構築と利用者間情報共有への展望. レコード・マネジメント, No.41, pp.8-18, Nov. 2000.
- [4] 小島篤博, 宮本貴朗, 富坂敏子, 小池利栄子, 石井敬三, 山野美贇子. 貴重文化資料デジタルアーカイブの Web による検索・共有システム

の構築. 電子情報通信学会技術報告, ET2001-31, pp.61-68, Jul. 2001.

- [5] 小島篤博, 青木茂樹, 宮本貴朗. 中尾佐助資料公開のための画像データベースの構築と運用. 公開シンポジウム「人文科学とデータベース」, Vol.16, pp.1-8, 2010.
- [6] 中尾佐助. 秘境ブータン (岩波現代文庫). 岩波書店, 2011.
- [7] 中尾佐助. ヒマラヤの花. 東京, 毎日新聞社, 1964.
- [8] 中尾佐助, 西岡京治. ブータンの花. 東京, 朝日新聞社, 1984.
- [9] “DSpace Official Site”. <http://www.dspace.org/>
- [10] “大阪府立大学学術情報リポジトリ OPERA”. <http://repository.osakafu-u.ac.jp/dspace/>
- [11] T.ヒース, C.バイツァー (武田英明監訳). Linded Data: Web をグローバルなデータ空間にする仕組み. 近代科学社, 2013.