

AIの進展に伴うメタデータに関する 図書館職員の変わる役割と変わらない役割

～機関リポジトリとデジタルアーカイブに関する京都大学の事例から～

京都大学附属図書館 研究支援課長
野中雄司

24

Slides

ジャパンリンクセンター 第12回 対話・共創の場
テーマ：AI時代におけるメタデータの再考
2026(令和8)年2月2日(月)
@JST東京本部別館 1階ホール(ハイブリッド)

北海道大学
附属図書館

13年

係員10年
係長3年

室蘭工業大学
附属図書館

3年

係長3年

東京大学
附属図書館

4年

係長4年

富山大学
附属図書館

3年

課長3年

京都大学
附属図書館

2年目

課長2年目



2025年9月
アレキサンドリア図書館@エジプト

現在の所掌

電子ジャーナル
データベース
の整備

デジタル
アーカイブ

図書館システム
含む
システム全般

オープンアクセス
機関リポジトリ



本日は
「デジタルアーカイブ」
「機関リポジトリ」
担当の立場から

本日のお話しの概要

【本会について（本セミナーウェブサイトより）】

AIが膨大な情報を直接解読・処理できるようになり、自然言語によるコンテンツの検索や、情報の抽出・要約、メタデータの自動生成も技術的に可能となってきました。この急速な変化に伴い、メタデータの役割も変化してきており、従来のように厳密なメタデータを作成する必要性そのものが問い直される状況も生まれています。

【本日のお話しの概要】

この状況において、現在図書館職員として、どんなことを考えながら、機関リポジトリ、デジタルアーカイブを運用しているのか、京都大学の現在及び計画中の事例を紹介しながら、考えてみます。

※ ちなみに私見多数込みです！（いわゆる機関代表ではなく）

シナリオ

1

AIの進展に伴って考えたこと

2

図書館職員の役割

3

京都大学における現状の取り組み

4

図書館職員の役割（再）

1

AIの進展に伴って考えたこと

知識を取得するための壁が乗り越えやすくなってきた？

7

1990年代終盤から、インターネットにより知識取得の**距離的な壁**が乗り越えやすくなった。（Eメール、電子ジャーナル等）



さらに生成AIの登場により、さらにいろいろな壁が低くなりつつある？

- 言語の壁
- 年齢や経験の壁
- 知識量の壁
- メディアの壁
 - テキスト、動画、音声等



画像は生成AI（Nano Banana）で生成

生成AIの学習データとしての学術図書や学術論文等

知識を得るため、AIにより壁が低くなっていくのはいいこと！

壁を低くするため、学術図書や学術論文等は、**生成AIの学習データとして**極めて重要な学習データだろう（現状一部のみ？
arXiv, PMC等オープンアクセスのもの中心？）



AI企業と大手出版社の契約が始まっている？



知識の囲い込み？（自社や特定AI会社のための学習）

知識の囲い込み？

(AI会社に学習用データとして販売しているケース)



出版社がコンテンツをAI会社にAI学習用データとして販売？

図書館界、図書館情報学

CA-R
カレントアウェアネス-R

CA-E
カレントアウェアネス-E

CA
カレントアウェアネス

[ホーム](#) » [カレントアウェアネス-R](#)

Taylor & Francis社、AI活用に向けた提携の一環として同社のコンテンツやデータへのアクセスをMicrosoft社に提供（記事紹介）

© 2024年08月23日

Taylor & Francis社が、AI活用に向けた提携の一環として同社のコンテンツやデータへのアクセスをMicrosoft社に提供したことが複数のメディアで報じられています。

<https://current.ndl.go.jp/car/224458>

THE BOOKSELLER

All in for the book trade since 1858

[COMMENT](#) [BESTSELLERS](#) [BOOKS](#) [SPOTLIGHT](#) [EVENTS](#) [JOBS](#)

約123億円！

Taylor & Francis set to make £58m from AI in 2024 as it reveals second partnership

NEWS JUL 25, 2024 BY MATILDA BATTERSBY

<https://www.thebookseller.com/news/taylor-francis-set-to-make-58m-from-ai-in-2024-as-it-reveals-second-partnership>

(出版社が自社コンテンツは自社のみで利用するケース)



HOME

囲い込みは抄録までも？！

Publishers close access to scholarly content such as abstracts due to AI incentives

(27 Sept 2025) The attached article discusses the recent for scholarly articles from major publishers like Elsevier aggregators such as OpenAlex. Over recent years, AI-po on open abstracts for effective research discovery.

2024年11月以降、これらの出版社は非オープンアクセス論文の抄録へのアクセスをますます制限するようになり、OpenAlexにおける2022年から2024年のエルゼビア論文の抄録のカバー率は、以前の約80%から40%未満、さらには22.5%まで大幅に低下しました。この変化は、出版社がAIのトレーニングと発見における抄録の価値を認識し、ライセンス契約を交渉しているためであり、これらのAIツールがこれまで享受していた「ただ乗り」が事実上終了しています。(Google翻訳)

However, since November 2024, these publishers have increasingly restricted access to abstracts for non-open access articles, drastically lowering abstract coverage—from around 80% previously to below 40% and as low as 22.5% for Elsevier articles from 2022-2024 in OpenAlex. This change is due to publishers recognizing the value of abstracts for AI training and discovery and negotiating licensing deals, effectively ending the previously “free ride” these AI tools enjoyed. The implications are wide-ranging: AI research search tools will have reduced effectiveness. normal keyword searches in open aggregators will miss many



二重取りから三重取りへ？

学術出版の世界では、今後、三重取りになっていく？

現状（学術雑誌の観点から）

- 「①購読料」と「②APC（論文投稿料）」の二重取り



近い将来（もう始まっている？）

- 上記に加え「③AI学習データ料」の三重取り

注）一部の商業出版社等

「知識の囲い込み」の観点からもオープン化推進

生成AIに対し、有料コンテンツ（売れる優良学習データ）
保有者はどのような判断をしていくのか？
（学術書、学術論文、新聞記事、等々）

この状況下で、オープン化は生成AIの学習データの観点から
も、さらに重要かつ議論の対象になっていくと思われる。

...と思われるが、難しいことは抜きにして？図書館職員とし
ては、壁を低くしていくために、まずはオープン化に尽力
したい

2

図書館職員の役割

話者私見です

私たち（図書館職員）ができること、役割

このような生成AIの状況を注視しつつも、私たちができることは？

① 1件1件コツコツとオープンにする

やっぱりオープン化は重要。私たちのできる範囲で、できる量をコツコツとオープンにしていくこと。

（特に世界中で私たちにしかできない、所蔵古典籍や、出版している紀要、学位論文など）

② 流通にのせ世界中の人に届ける

加えて、オープンにした成果物を、きちんと整理し世界中の人に届けること。（メタデータ（書誌情報など）を整え、世界中の人がみつけやすいようにすることも含む）

生成AIの学習データとしての視点から

さらに意識している（し始めた？）こと

1. 本文テキスト（抄録等含む）のオープン化

- 本文テキストのオープン化は、生成AIの学習データとしても意義があるだろうことも意識したい
- 特に言語の壁も低くなっており、日本語（ローカル言語）文献の重要性がさらに高まるのでは。

2. 書誌情報等のメタデータの整備と流通

- 人間によるメタデータの整備は、ハルシネーションを減らす観点からも重要ではないか（これまで私たちの重要な仕事であったが、改めてその重要性も高まっているのでは）

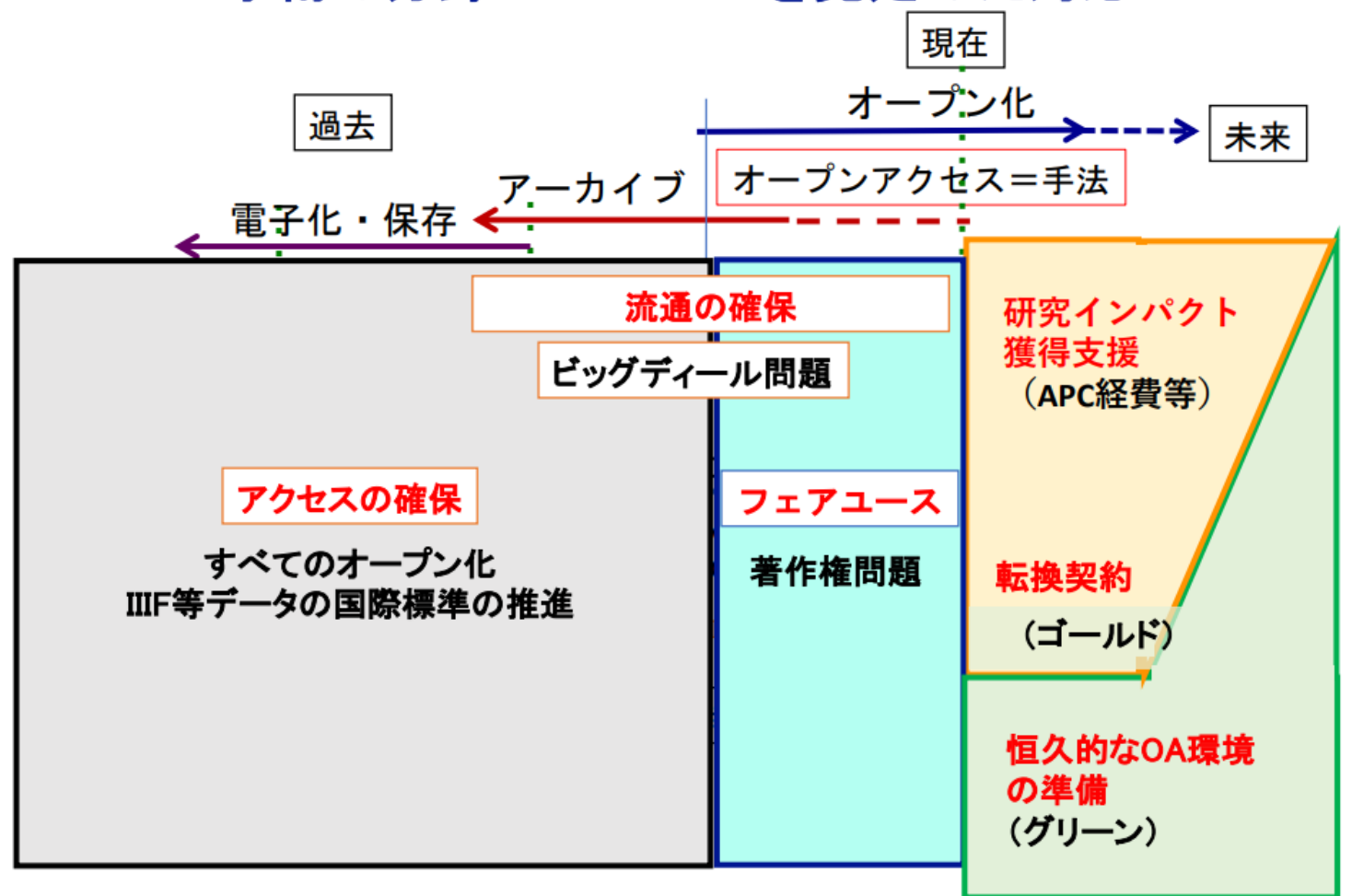
3

京都大学における現状の取り組み

学術のオープン化に向けた図書館の取り組み

学術情報の閲覧可否状況

学術の分野・フェーズを見定めた対応



京都大学では、灰色と緑色部分については、主に以下の2サービスで展開中

京都大学学術情報リポジトリ
(オープン化のプラットフォーム)
【主に緑部分】



京都大学貴重資料アーカイブ
(アーカイブのプラットフォーム)
【主に灰部分】



(青色と黄色については本日は省略)



京都大学学術情報リポジトリ (KURENAI) では、京都大学で日々創造される研究・教育成果（学術雑誌掲載論文、学位論文、紀要論文など）をWeb上で公開しています。世界的に卓越した知的成果を社会へ還元することを目的として、2006年から図書館機構が運営している事業です。

京都大学学術情報リポジトリ **KURENAI** 紅 総公開資料数
247,194件

資料種別	件数
紀要論文 Departmental Bulletin Paper	129,259
学術雑誌論文 Journal Article	46,373
学位論文 Thesis or Dissertation	31,709
その他 Others	21,596
論文 Article	9,914

KURENAIによるオープン化（主なコンテンツ）

- 学術雑誌論文：46,373件
(2025年度：2,207件)
- 紀要論文：129,259件
(2025年度：1,133件)
- 学位論文：31,709件
(2025年度：649件)
- 研究データ：338件
(2025年度：26件)

2025年度は2025年10月現在の値

貴重資料デジタルアーカイブ


日本語 English

京都大学貴重資料デジタルアーカイブ
 コレクション 検索 お知らせ サイトについて 京都大学図書館機構



肥後国海中の怪、アマビエの画像をご利用いただく際には
肥後国海中の怪、アマビエの画像をたくさんの方々にご利用いただき、ありがとうございます。オープンアクセスデータは、誰もが自由に使うことのできる大切な共有財産です。ご利用の際に知っておいていただきたいことをまとめてみましたので、データを使う人、制作物を見る人、皆が気持ちよく利用できるような、ご協力をお願いいたします。

● 説明サイトを見る



貴重資料デジタル化プロジェクトへのご支援を募集します
京都大学が所蔵する貴重な古典籍資料のデジタル化・公開を進めるため、「京都大学貴重資料デジタルアーカイブ基金」を設置しました。詳細はこちらをご覧ください。



画像を自由に二次利用できる資料所蔵図書館・室を拡大しました
インターネット上で公開している京都大学附属図書館、吉田南総合図書館、法学部図書室、経済学研究科・経済学部図書室、理学研究科各図書室及び理学部中央図書室、基礎物理学研究所、総合博物館所蔵資料の電子化画像は、利用申請・利用料の支払手続きをすることなく、自由に利用することができます。詳細は「コンテンツの二次利用について」をご覧ください。

お知らせ

2025-10-07 京都大学アカデミックデイ2025に参加しました
 2025-09-22 京都大学アカデミックデイ2025に参加します（9/27）
 2025-04-24 利用報告のためのフォームを設置しました
 2025-03-27 総合博物館が所蔵する「教王護国寺文書」より310点を公開しました
 2025-03-27 画像の不具合について

● すべてのお知らせを見る

ピックアップ



国宝・今昔物語集
(鈴鹿本)



重要文化財



時代の記録



彩りの挿絵



地図でみる日本、世界



京都大学所蔵資料
でみる博物学の時代



京都大学所蔵資料
でたどる文学史年表



【学内者対象】公開
したい画像をお持ちの方へ



展示会・企画展・
研究成果

貴重資料デジタルアーカイブ によるアーカイブ公開

- 図書館・室所蔵の古典籍を中心としたアーカイブと公開
- 25,905 タイトル
- 2,155,567 画像
- 国宝・重要文化財等含む

2024年度末現在の値

生成AIの学習データとしても意識したKURENAIでの取り組み

1. 本文テキスト（抄録等含む）のオープン化

- 古い紀要論文などスキャンしたPDFを掲載する場合は、必ずOCRをかけて透明テキストを埋め込む（従来からGoogle等でヒットするように実施しているが、さらに重要に？）
- 紀要論文（京都大学発オリジナルコンテンツ）本文について、PDFだけではなく、XML（HTML）での提供を検討中（プラットフォームの変更（例：DSpace→OJS）

2. 書誌情報等のメタデータの整備と流通

- 多様なデータベース（IRDB, Google Scholar, PubMed, OpenAlex等）から採録されるよう、さらに積極的に対応
- 紀要論文のDOIについて、JaLC DOI付与からCrossref DOI付与へ変更（2026年4月から予定）
- 研究データについては、（従来より）全てDataCite DOIを付与

1. 本文テキスト（抄録等含む）のオープン化

- NDL古典籍OCRを利用し、全ての画像の本文データ化・公開を計画中（2026年度内目標）

2. 書誌情報等のメタデータの整備と流通

- 全てのコンテンツ（書誌単位）にJaLC DOI付与を行う（2026年4月からを目標）
- 公開を優先し、目録はラフに作っているが、きちんとした遡及目録作成も重要（少しずつでも継続的に実施していく）

4

図書館職員の役割（再）

話者私見です

これまでも、これからも

（これまで（も）） 知識を伝播する役割

私たちは世界中で協力して、1件1件目録を作成し、使いやすいように棚に並べて、知りたい人が知りたいことにアクセスできるような環境を構築する仕事をしてきた。

（これから（も）） 大学発の研究資源・成果を流通させる役割

加えて、近年大学の研究資源や研究成果を公開する役割も。
生成AIの意識はするが、基本は1件1件オープン化や研究成果発信を行い、整理、流通させ、世界中の知りたい人が知りたいことにアクセスできるようにしていく。

これらは**私たちの最も得意**とすることであり、かつ**私たちにしかできないこと**ではないか。

個人的まとめ

生成AIにより世界が変わろうと、私たちの基本的な役割は変わらない。

私たちは変わらずそれに貢献できるはずであり、むしろその重要性は増している。

戦争、貧困などの世界的諸課題を解決するための基盤を構築し、社会や科学の発展に貢献していく。