人間文化研究情報資源共有化研究会報告集 9

Proceedings of the Study on Information Resources of the Human Science Vol.9

令和元（2019）年 6 月

June 2019

大学共同利用機関法人　人間文化研究機構

総合情報発信センター

高度連携情報技術委員会

目　　次

人間文化研究情報資源共有化研究会報告集９の刊行にあたって　　永村　眞 …… 3

第14回研究会報告集

第14回研究会プログラム …… 5

講演

「nihuINT の「後」とは何か」
大内英範（人間文化研究機構本部）
要旨 …… 7
スライド …… 8

「関西大学 KU-ORCAS のデジタルアーカイブ構築とそのデータ利活用に向けて」
菊池信彦（関西大学アジア・オープン・リサーチセンター）
要旨 …… 13
スライド …… 14

「歴史資料の共有化とデータ連携・利活用」
榎原雅治（東京大学史料編纂所）
要旨 …… 27
スライド …… 28

「商用による情報提供と公的な情報提供の関係、未来」
田中政司（株式会社ネットアドバンス）
要旨 …… 35
スライド …… 36

パネルディスカッション …… 45
利用者のニーズをどう汲み取ってきたか
データベースを作る側としての研究者はどうあるべきか
データを作るためのデータ
人間文化研究機構への要望

人間文化研究情報共有化研究会報告書９の刊行にあたって

大学共同利用機関としての人文学系の６研究機関（国立歴史民俗博物館、国文学研究資料館、国立国語研究所、国際日本文化研究センター、総合地球環境研究所、国立民族学博物館）から構成される人間文化研究機構は、各機関の研究と情報化の事業を基礎として、機関の枠を越えた連携をはかり、国内外への情報発信などを行う役割をになってきました。機構として独自の研究事業をになうことはできませんが、６研究機関の研究業務を支援し、その高度化を実現するとともに、成果の広範な発信などを、果たすべき任務としてきました。

この６研究機関の連携という事業を推進するため、平成20年（2008）に研究資源共有化事業委員会が生まれ、さらに平成28年からは高度連携情報技術委員会がその実務を継承しています。また平成21年から人間文化研究情報資源共有化研究会が発足して、毎年新たなテーマのもと研究発表と討論が開催され、平成30年度には「人間文化研究機構が持つデータの役割とその未来」という課題で、大阪大学中之島センターを会場に研究会がもたれました。本年度のテーマは、人間文化研究機構の情報処理事業の将来像を見すえるための、従来の事業の総括とともに、他機関における類似した事業との関係について、本報告書に収められた様々な提言と議論が重ねられました。

さて人文学系の研究は、歴史・民族・民俗・言語・文化・環境など多様な分野に分かれ、その内実は文字情報・画像情報・動画情報・音声情報・空間情報など多彩な広がりがあります。そこで縦割りの研究分野にとらわれぬ、つまり分野を超えた情報の共有化を実現するために開発された処理システムこそがnihuINTであり、本システムを稼働することにより各機関の多様な形式のデータベースを連携させ、大規模なデータ群の横断検索が可能となりました。まさにnihuINTは、情報事業という側面で人間文化研究機構の役割を象徴するシステムと言うことができましょう。

今後の人間文化研究機構は、その本来の役割としてある大学共同利用機関としての、望ましい機能のもとで、６研究機関のみならず、広く諸大学の情報事業にも関わりをもち、その効率的な処理・発信の事業を担う必要があります。そして４年後に始まる第四期の情報処理事業は、各研究分野の要請に応える取り組みが求められます。人間文化研究機構がもつ情報の連携・発信のためのシステムは、各機関が開発に関わり蓄積してきた貴重な成果であり、このシステムがもつ多彩な機能を柱として、次の時代に貢献できる新たな枠組みを策定し、そのもとで情報事業の新たな展開を目指したいと思います。

本研究会の課題は新たな研究分野の創成とともに、現状を踏まえて将来にわたる情報処理事業の展開を如何に実現するか検討することであり、その提言と討論のなかに次の時代における人文学情報のあり方の予見を期待いたします。

2019年６月

大学共同利用機関法人　人間文化研究機構
客員教授　永　村　　眞

第14回人間文化研究情報資源共有化研究会

「人間文化研究機構が持つデータの役割とその未来」

人間文化研究機構において第３期のnihuINTが公開されて、２年が経過した。人文学の研究データを取り巻く状況は、データ基盤の整備状況や、オープンデータの動向、人文学の「危機」といった議論を含め、大きく変化しつつある。そのような状況の中で、人間文化研究機構が持つべきデータ、持つべきシステムとはどのようなものであるべきか、次のステップのための議論をさらに深めなければならない。本研究会では、nihuINTのありようそのものから含め、「ゼロベース」での議論を行い、人間文化研究機構のみならず、人文学が持つデータの形式・保持方法・利用方法などについて、様々な見地から検討したい。

○　**日　時**　平成31年２月15日㈮　13時30分～16時30分（13：00開場）

○　**会　場**　大阪大学中之島センター（507講義室）
大阪市北区中之島４－３-53

○　**プログラム**

13：30　趣旨説明　人間文化研究機構　永村　眞

13：40　「nihuINTの「後」とは何か」　人間文化研究機構　大内　英範

14：05　「関西大学KU-ORCASのデジタルアーカイブ構築とそのデータ利活用に向けて」　関西大学アジア・オープン・リサーチセンター　菊池　信彦

14：30　休憩

14：40　「歴史資料の共有化とデータ連携・利活用」東京大学史料編纂所　榎原　雅治

15：05　「商用による情報提供と公的な情報提供の関係、未来」
株式会社ネットアドバンス　田中　政司

15：30　休憩

15：40　ディスカッション「人文学データの今後と人間文化研究機構が目指す情報基盤」
司会：国際日本文化研究センター　関野　樹

16：30　閉会

○ **主　催**　大学共同利用機関法人 人間文化研究機構 総合情報発信センター
高度連携情報技術委員会

○ **連絡先**

人間文化研究機構本部 情報発信センター係
〒105-0001　東京都港区虎ノ門４－３－13　ヒューリック神谷町ビル２階
（TEL）03-6402-9234（FAX）03-6402-9240　（Mail）cip-office@nihu.jp

nihuINT の「後」とは何か

人間文化研究機構本部 総合情報発信センター　大　内　英　範

nihuINT（nihu INTegrated retrieval system、https://int.nihu.jp）は、多種多様なデータ構造を持つ人文科学のデータベースを一元的・網羅的に検索できる統合検索システムとして、2008年に運用を開始した。機構の6機関および機構が運用する地域研究拠点のDBのほか、外部機関のDBとの連携も実現し、第2期終了時点（2015年度末）で検索対象は172DBに拡大していた。多種多様かつ大量のデータの横断検索は、「思いがけない発見」が期待され、それを支援する仕組みも搭載されていた。

2016年度末にリプレイスされた3代目の現行nihuINTは、スマートデバイスへの対応のほか、Linked Dataへの対応が重要な新機能である。2021年度までの現行nihuINTの期間中に、可能な限り既存データのRDF化を進めること、そして、それらのデータの入口となるような基盤データの整備を着々と進めてゆくことが重要だと認識している。

そして、その「後」つまり第4期のnihuINTであるが、大量のデータを背景とした「思いがけない発見」の支援から、研究に必要な基盤データの整備・拡充に重心を移すことを考えている。それに伴い、現行機能の取捨選択および積極的な外部サービス利用を行なってスリム化し、APIの整備による多様なデータ利用を可能とするものとしたい。その一端を報告する。

nihuINTの「後」とは何か

2019.02.15

人間文化研究機構・資源共有化研究会@大阪大学中之島センター
人間文化研究機構本部・総合情報発信センター　大内英範

1

はじめに

- 2008年　統合検索システム運用開始（2011年6月～nihuINT）

- 2016年度～　第3期中期目標期間
- 2016年度末　現行nihuINT運用開始
- 2018年度末　←イマココ

- 2022年度～　第4期中期目標期間

2

nihuINTとは

- nihuINT：nihu INTegrated retrieval system、https://int.nihu.jp

- 2008年4月の公開当時は「統合検索システム」
  - 機構内5機関のデータベースの横断検索
  - 2008年12月、nihuONEシステム公開

- 2010年7月、国立国会図書館PORTAとの双方向連携開始
    - この時点で119DB
  - 2010年9月、時空間解析ツール（GT-Map、GT-Time）のサービス開始
  - 2010年12月、国立国語学研究所のDBも参加
  - 2011年6月からnihuINT、URLをhttp://int.nihu.jp/ に
  - 2015年、京都大学地域研究統合情報センターのDBとの双方向連携開始

3

nihuINTとは(2)

- 2017年3月、現行nihuINT運用開始
  - 機構内外の172DB
  - 特徴
    - 目的志向型検索
    - スマートデバイス対応
    - SNS連携機能
    - Linked Dataへの対応
  - 検索速度の向上
  - 機構内機関の機関リポジトリとの連携
  - 検索数は月間6万件前後で推移

4

nihuINTの役割（私見含め）

- 機構内5機関（当初）の多様なDBを、その所在や各操作方法などを意識することなく、まとめて1回の操作で検索するためのシステムとして作られた。その後対象を拡大し、人文学研究情報のハブ的存在へ
- 特に2期システムでは、検索結果に関連するデータを示すファセットナビゲーション機能があり、多種多様かつ大量のデータを背景とした「思いがけない発見」を支援するシステムだった
  - 本屋さんの本棚ぽい?
  - 検索速度・表示速度に大きな影響を及ぼしていたので現行3期システムには継承されず
  - ※3期システムにも「類似資料を開く」機能あり

5

nihuINTの役割（私見含め）

- 3期システムでは、情報の共有・連携を進める機能を強化。情報をつなぐ役割
- スマートデバイス対応／SNS連携機能
  - 手軽に検索、twitter等に送信して情報共有・話題提供
- Linked Dataへの対応
  - nihuINT-LD（実験中）では、RDFに変換したnihuINTのデータの検索、内外のデータとの連携
  - 3期中（来年度目標）に一部公開したい

6

nihuINTの「後」

- これまでnihuINTは対象拡大・機能追加／強化等を行ないながら、基本的なシステムをほぼそのまま継承してきた
- 4期はどうするのか

- さまざまな状況・環境の変化
  - 2022年度から4機構と総研大が運営統合
    - 新設する一般社団法人に5法人のまま参画
    - 予算や人員、調達、情報セキュリティー、社会貢献窓口など多くの業務を一元化し、効率化や柔軟な資源配分を追求する
    - →ますます交付金減が見込まれる
  - 2018年度末に内閣府・国会図書館の運営する「ジャパンサーチ」が一般試験公開（予定）
    - 国立文化財機構、国立公文書館、国立科学博物館、早稲田大学、NHK等々が参加
    - 人間文化研究機構からは、nihuINT搭載データを中心に50以上のDBを提供（予定）

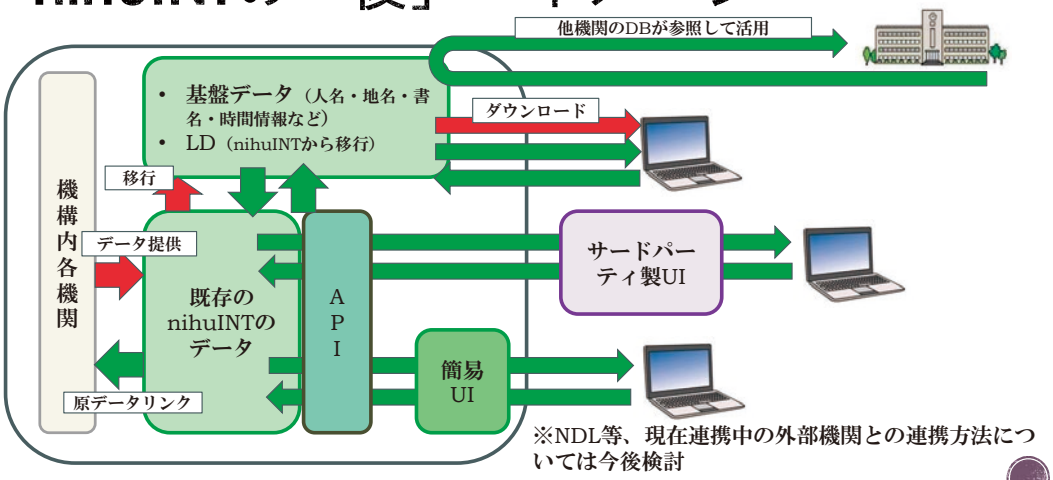
7

nihuINTの「後」(2)

- これまでの、「データの探索」・「思いがけない発見」の支援という役割から、人文科学研究のための基盤データをしっかり整備して提供するという役割へシフトする
  - 基盤データ：人名、地名、書名、時間情報など
  - 既存のnihuINTのデータは引き継ぐが、UIは簡素化し、代わりにAPIを整備する
  - 既存のnihuINTのデータをなるべくLDに移行する
  - 利用できる外部サービスはなるべく利用する

8

nihuINTの「後」・イメージ



※NDL等、現在連携中の外部機関との連携方法については今後検討

9

おわりに

- 運用開始から10年以上経過したnihuINTの「後」
- データ、API、UI
- 既存のnihuINTのデータは継承しつつ、機能をかなり絞り込むかわりに、APIを整備・提供することで、何をどのように探すのかは、ユーザに委ねる
  - 分野ごとUIなどどんどん作ってもらう
- さまざまな個別データの入口になるような基盤データの整備と、既存データのLD移行を進める
  - ダウンロード、リンク、SPARQLなどでさまざまに利活用してもらう
- 3期残り3年
  - 「後」を見据えたさまざまな作業、実験
- 人文学の人文学による人文学のためのデータ整備・システム構築を

10

関西大学KU-ORCASのデジタルアーカイブ構築とそのデータ利活用に向けて

関西大学アジア・オープン・リサーチセンター　菊　池　信　彦

本報告では、平成29年度文部科学省私立大学研究ブランディング事業に採択された関西大学アジア・オープン・リサーチセンター（KU-ORCAS）について紹介する。具体的には、その事業の中心にあるデジタルアーカイブについて、登録対象となる資料群や３つのコンセプト、そしてそれらを包含するオープンプラットフォームにも言及する。また、その概要紹介を踏まえたうえで、今後の研究活動のなかで生み出されるであろう研究データについても、その管理や利活用に関する課題を指摘する。そして、KU-ORCASのプロジェクトを踏まえつつ、これからのnihuINTと人間文化研究機構への期待についても述べることで、日本のデジタルヒューマニティーズの定着と活性化を図るものとしたい。

関西大学KU-ORCASのデジタルアーカイブ構築と
そのデータ利活用に向けて

菊池信彦
（関西大学アジア・オープン・リサーチセンター）

第14回人間文化研究情報資源共有化研究会
日時：平成31年2月15日（金）
会場：大阪大学中之島センター

KANSAI UNIVERSITY

1

構成

本日の話の流れ

1. nihuINTを使ったことがありません
2. ハブを目指しているKU-ORCAS
3. KU-ORCASにおける研究データ
4. 機構への問題提起

KANSAI UNIVERSITY

2

1. nihuINTを使ったことがありません

KANSAI UNIVERSITY

3

1.nihuINTを使ったことがありません

- そもそもスペイン近現代史が専門だからではありますが、
- 図書館員の経験では、nihuINTが何を検索して何を検索できていないのかが分かりづらいから
- NDLサーチでもnihuINTは検索可能であり、また、NDLサーチの方が検索範囲が広い

NDLサーチへデータを提供するアグリゲーション
という認識

KANSAI UNIVERSITY

4

5

2. ハブを目指している KU-ORCAS

KANSAI UNIVERSITY

6

KU-ORCASとは

KU-ORCAS＝関西大学アジア・オープン・リサーチセンター（Kansai University Open Research Center for Asian Studies）

平成29年度文部科学省私立大学研究ブランディング事業に選定（5年間で現在2年目）

東アジア文化研究のためのデジタルアーカイブの構築と研究、それを通じた研究のブランディング

“東アジア研究といえば関大
関大といえば東アジア研究”

DRAW YOUR ASIA
KU-ORCAS
関西大学
アジア・オープン・リサーチセンター
KANSAI UNIVERSITY OPEN RESEARCH CENTER for ASIAN STUDIES

KANSAI UNIVERSITY

7

そのために、何を？どのように？

対象	コンセプト	方法
• 関大所蔵資料 • バチカン図書館のデジタルアーカイブへの協力	• 3つのオープン化	• オープンプラットフォーム

デジタルアーカイブ

KANSAI UNIVERSITY

8

2．ハブを目指しているKU-ORCAS

「関西大学の所蔵資料」とは

飛鳥等の考古学研究資料

東西言語接触に関わる資料群

泊園文庫・大坂画壇コレクション

他にも、
・東洋音楽資料
・日本占領下の映画パンフ
・朝鮮の碑石
・古代朝鮮の王陵・寺院
など

図書館資料や研究者個人蔵書・収集資料

KANSAI UNIVERSITY

9

2．ハブを目指しているKU-ORCAS

①研究リソースのオープン化
- IIIFへの対応（IIIFコンソーシアム加盟）
- 画像・メタデータのCCライセンスの付与

International Image Interoperability Framework

PUBLIC DOMAIN

②研究グループのオープン化
- クラウドソーシングによる翻刻プロジェクト
- 教育リソースの配信

YouTube

③研究ノウハウのオープン化
- DAの構築やDHにおけるデータ利活用の方法を共有

④研究成果のオープン化
- 論文や講演会資料等のオープンアクセス
- 研究データの公開・利活用推進

現在策定準備中

KANSAI UNIVERSITY

10

2．ハブを目指しているKU-ORCAS

IIIF International Image Interoperability Framework に対応した東アジア関係資料の統合検索ポータル

KANSAI UNIVERSITY

11

KU-ORCASの国際学術ネットワークと それをベースにしたIIIF対応東アジア関係資料の統合検索

IIIFコンソーシアム会員メンバーに対して、協力を打診

1. 大英図書館（英国）
2. ルーヴェン・カトリック大学（ベルギー）
3. フランス国立図書館（フランス）
4. バチカン図書館（バチカン市国）
5. ローマ大学（イタリア）
6. 北京外国語大学（中国）
7. 復旦大学（中国）
8. 浙江大学（中国）
9. 高麗大学校（韓国）
10. 国立中央研究院（台湾）
11. 国立台湾大学（台湾）
12. プリンストン大学（米国）
13. ハーバード大学（米国）

KANSAI UNIVERSITY

12

バチカンとの連携

- バチカン図書館所蔵資料のうち、日本関係15点のデジタル化を依頼。その後、目録作成に協力するという流れ
- クラウドファンディングの導入も検討中

KANSAI UNIVERSITY

13

オープンプラットフォームによって、
IIIF対応東アジア研究資料を一元検索し、
東アジアDHのノウハウや研究情報を集約することで
アグリゲーション機能を高め
東アジア文化研究のハブ組織を目指す

↓

「東アジア研究なら関大KU-ORCASだ」
とみなされるように

KANSAI UNIVERSITY

14

３．KU-ORCASにおける研究データ

KANSAI UNIVERSITY

15

３．KU-ORCASにおける研究データ

- KU－ORCASの研究ユニット（一部）

①東西言語接触
②大阪漢学文化（泊園書院）や近世大坂画壇
③明日香・難波津の古墳・考古学関係
④中国古代木簡の字体分析
⑤東アジア碑石、金石拓本、古代朝鮮の王陵等
⑥日本植民地下の映画や劇場文化研究
⑦ユーラシアにおける文化交渉　…

→ 各研究ユニット／研究者の活動から
デジタルアーカイブが作成され
研究データが蓄積される

KANSAI UNIVERSITY

16

3．KU-ORCASにおける研究データ

想定される研究データ群

- 東西言語接触　⇒　非構造化テキストデータ
- 泊園書院・大坂画壇　⇒　人物ネットワーク情報
- 古墳・考古学関係　⇒　地理情報
- 木簡の字体分析　⇒　文字データセット
- 碑石データベース　⇒　写真とそのメタデータ

⇒性質の異なるものがバラバラに登場することに

もちろん
公開できるものは利用しやすい形で公開が大前提

KANSAI UNIVERSITY

17

3．KU-ORCASにおける研究データ

- 課題は

①対象：誰に向けて

②方法：どこで、どのようにメタデータを付けて、どの程度の情報を

④永続性：提供し続けることができるような体制をどのように作るのか

⇒　各機関の課題として「のみ」悩むべきか？
むしろ関連機関共通の悩みでは？

KANSAI UNIVERSITY

18

４．ハブを期待したい人間文化研究機構への問題提起

KANSAI UNIVERSITY

19

本日の問題設定を振り返る

「人間文化研究機構が持つべきデータ、持つべきシステムとはどのようなものであるべきか、次のステップのための議論をさらに深めなければならない。本研究会では、nihuINTのありようそのものから含め、『ゼロベース』での議論を行い、人間文化研究機構のみならず、人文学が持つデータの形式・保持方法・利用方法などについて、様々な見地から検討したい。」（研究会趣旨文より）

⇒研究ライフサイクルから生まれた研究データ管理とは切り離して考えるべき

KANSAI UNIVERSITY

20

4．機構への問題提起

トップダウンから機構に期待される役割

第4期中期目標期間における大学共同利用機関の在り方について（審議のまとめ）　2018年12月14日

- 「大学共同利用機関法人は，時代の要請に応じて，新たな学問分野の創出に戦略的に取り組むことが必要であり，設置する大学共同利用機関について，各研究分野の動向，大学の研究者のニーズ，将来性等を踏まえ，その在り方を検討することが必要」（pp.8-9）
- 「『大学共同利用機関として備えるべき要件』については，主に以下のような内容が考えられるところであり，今後，文部科学省において，科学技術・学術審議会の意見を聴き，法令等において具体的に定めることが必要である。
  ・開かれた運営体制の下，各研究分野の研究者コミュニティ全体の意見を取り入れて運営されていること
  ・各研究分野に関わる大学や研究者コミュニティ全体を先導し，最先端の研究を行う中核的な学術研究拠点であること
  ・国際的な学術研究拠点として，各研究分野における我が国の窓口としての機能を果たしていること
  ・個々の大学では整備・運用が困難な最先端の大型装置や貴重な学術データ等の研究資源を保有し，これらを全国的な視点に立って共同利用・共同研究に供していること 」（p.9）

http://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/010/toushin/1412585.htm

新たな学問分野の創出と各研究分野への精通

KANSAI UNIVERSITY

21

4．機構への問題提起

「人文学・社会科学の振興に向けて（審議のまとめ）」平成30年12月14日
科学技術・学術審議会学術分科会人文学・社会科学振興の在り方に関するワーキンググループ

- 「人文学・社会科学分野ではデジタル化に未着手の数多くの領域があるが、デジタル人文学などの取組が進みつつあるように、多様な研究データのデジタル化された利用環境が充実するとともにデータサイエンスを応用したデータ駆動型の新しい手法を取り入れることにより、これまで得られなかった学術的・社会的成果を人文学・社会科学研究が生み出すことに大きな期待が寄せられている。また、オープンサイエンスという視点から、研究の基盤となるデータの公開を通じた共同利用の促進に向けた動きも並行して進められている。」（p.4）

http://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/044/houkoku/1412891.htm

人文学の振興という文脈でDHへ期待が寄せられる現状

KANSAI UNIVERSITY

22

4．機構への問題提起

今日の趣旨と昨今の議論の流れから考えると、

「各研究分野」＝DHとみなしてよい

⇒機構と大学共同利用機関が

- 「DHのコミュニティ全体の意見を取り入れる」
- 「DHに関わる大学や研究者コミュニティを先導し…」
- 「DHにおける我が国の窓口として…」

国内DH拠点のそろい踏み

KANSAI UNIVERSITY

23

4．機構への問題提起

したがって、人間文化研究機構がボトムアップな観点からDHを行うだけではなく、

- 人間文化研究機構という一組織としての視点ではなく、国内DHのハブ組織としての視点に立ち、
- 研究データをDHで利用する場面を想定し、
- どのようなデータをどのように管理・利用すればよいのかについて、日本のDHの研究動向と研究者ニーズの調査を行ってみては？

人文学データ利活用のための検討土台となる

「研究データ」の作成を期待します

KANSAI UNIVERSITY

24

歴史資料の共有化とデータ連携・利活用

東京大学史料編纂所　榎　原　雅　治

人間文化機構の各機関と同様、国内外の大学や研究機関にも多くの歴史史料が所蔵されており、現在では各機関から画像をはじめとする史料情報が公開されている。本報告では、東京大学史料編纂所と人間文化機構、とりわけ歴史民俗博物館（歴博）の所蔵する史料情報の連携について考えてみたい。

史料編纂所と歴博の公開するデータベースの中には、全く同じ史料に関する情報が含まれている場合がある。機関としての全体的なデータ蓄積や紙媒体を含む双方の発信のあり方にはそれぞれメリットとデメリットがあり、相互に相補えば利便性が向上することは明らかである。しかし残念ながら、現時点では両者は全く無関係に発信されている。歴史の研究者としてはどのような連携を望むか、具体的な事例に即して提示したい。

もう一つ、現在、東京大学では、史料編纂所と地震研究所が連携して近世の日記史料から大小の地震動の記録を収集するプロジェクトを立ち上げている。近世には各藩、庄屋、商家などさまざまな階層で大量の日記が書かれたが、その存在の全国的な情報はどの機関も把握していない。人間文化研究機構の行っている「歴史文化資料保全の大学・共同利用機関ネットワーク事業」と連携した史料所在情報の収集に期待するところである。

第14回人間文化研究情報資源共有化研究会

歴史資料の共有化とデータ連携・利活用

東京大学史料編纂所　榎原雅治

本日の話題提供　史料編纂所と人間文化研究機構（特に歴史民俗博物館）とのデータ連携についての期待

人間文化研究機構の資源共有化事業において、当初追究されていたデータ連携の方法

nihuINTへの参加　⇒　データの一元的管理　さまざまな事情によって実現しなかった

・人文機構側の事情　　nihuINTに取り込むためには多大な時間と費用が必要

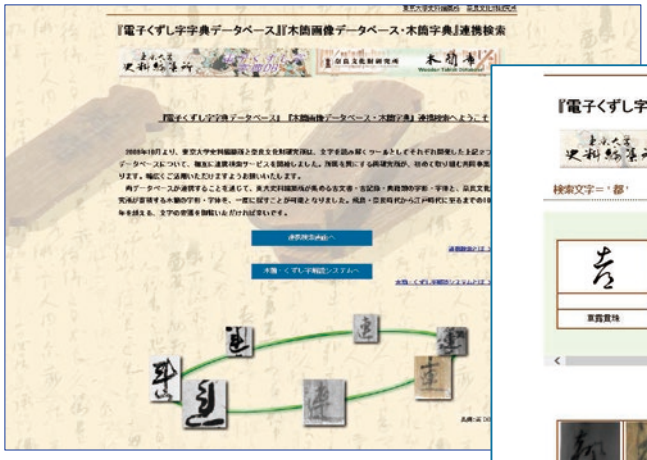
・史料編纂所側の事情　　所蔵史料情報は研究所の事業展開のための重要な資源
DBへのアクセス数は評価対象

歴史史料を所蔵する機関はいずれも同様の事情をかかえていると思われる

nihuINTへの参加機関を増やす形ではない方法でのデータ連携が必要

1

1　史料編纂所や人文機構ですでに実施されているデータ連携の事例



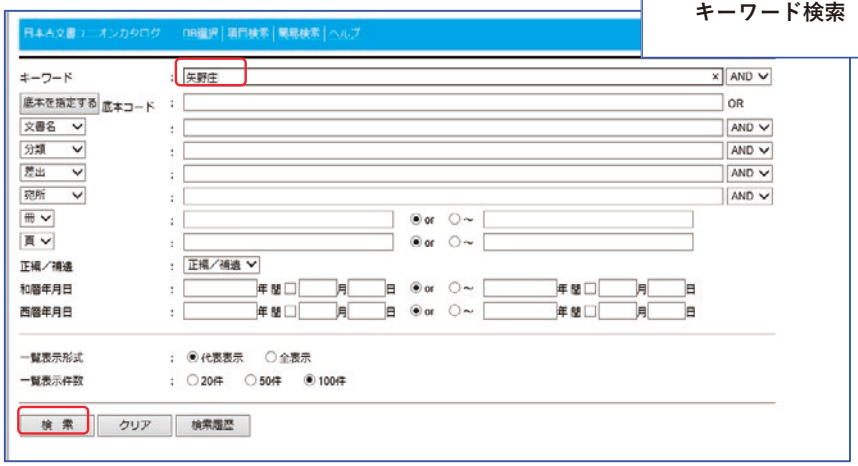
史料編纂所と奈良文化財研究所の連携検索

2

史料編纂所と京都学・歴彩館のDBリンク

編纂所《古文書ユニオンカタログ》

キーワード検索　⇒　検索結果表示



3

《古文書ユニオンカタログ》

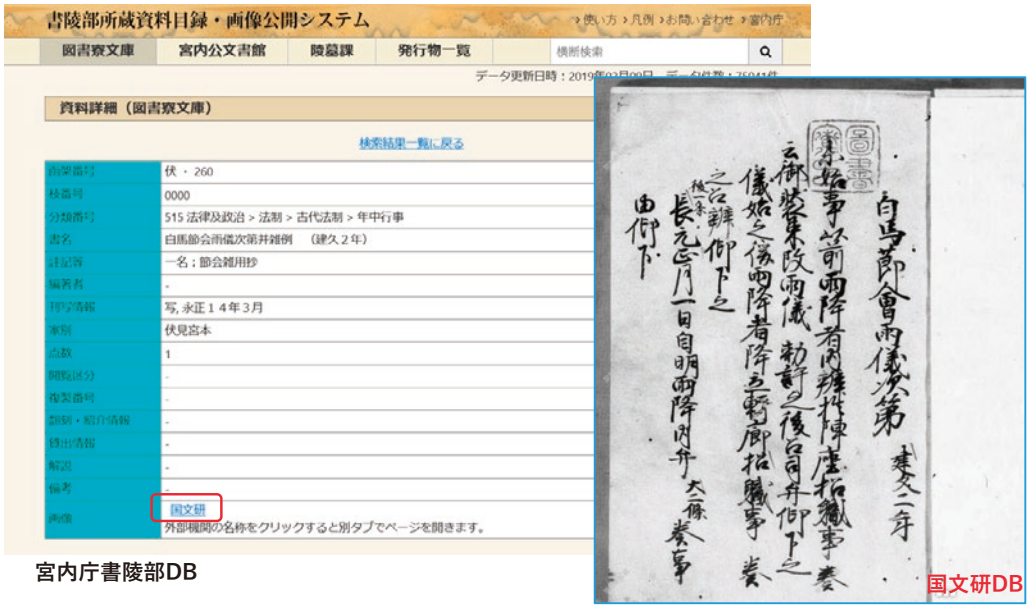
検索結果表示　⇒　詳細表示

⇒　京都学・歴彩館DBイメージ

4

宮内庁書陵部と国文学研究資料館のリンク

5



6

2　史料編纂所DBと歴博館蔵資料DB

研究の対象とする資料は非常に近い関係にある
同じコレクションに関する情報を公開（原本所蔵歴博　写真帳編纂所）

例：広橋家旧蔵記録文書典籍類（広橋家記録）　田中穣氏旧蔵典籍古文書

編纂所

1点目録あり

写真帳　広橋126冊、田中261冊
（閲覧室で公開）

目録以外の情報なし

歴　博

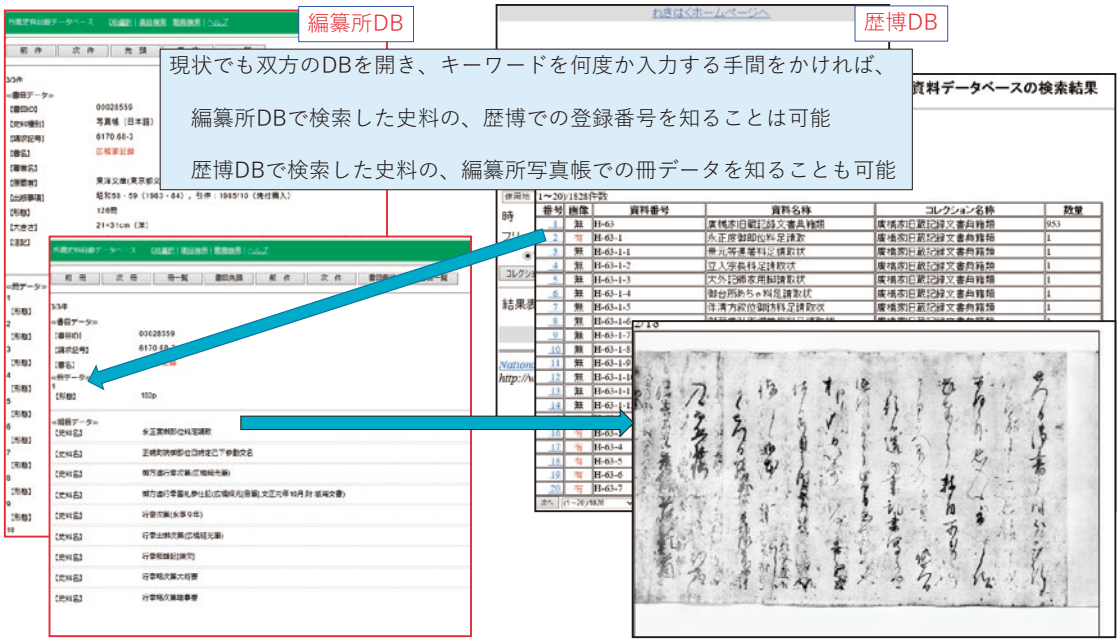
1点目録あり

画像あり（サイズが小さく判読には難）

備考に詳細情報あり

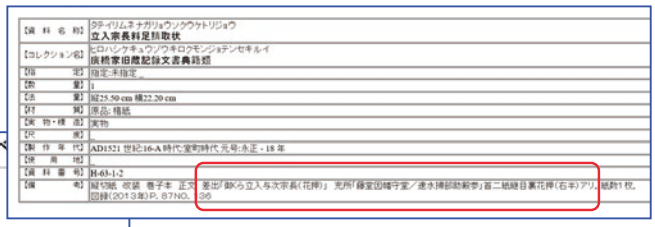
現状では無関係に公開・発信

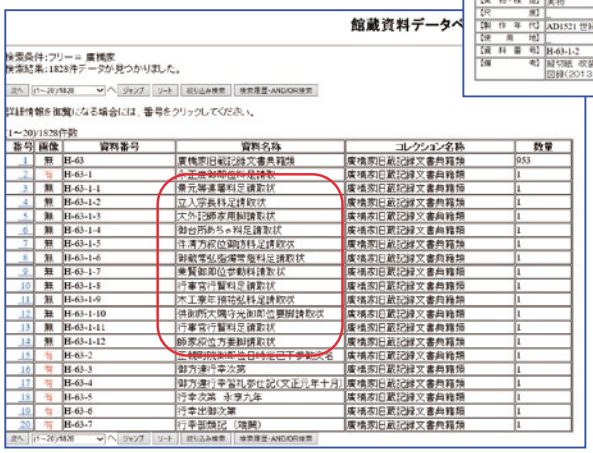
7



8

リンクにとどまらない相互補完的な情報流通が実現できれば、研究上の利便性は高まる





歴博DBの特性

編纂所DBがもっていない情報

編纂所DBでは検索できない史料細目、人名、役職名の情報が含まれている

編纂所DBから検索することはできないか

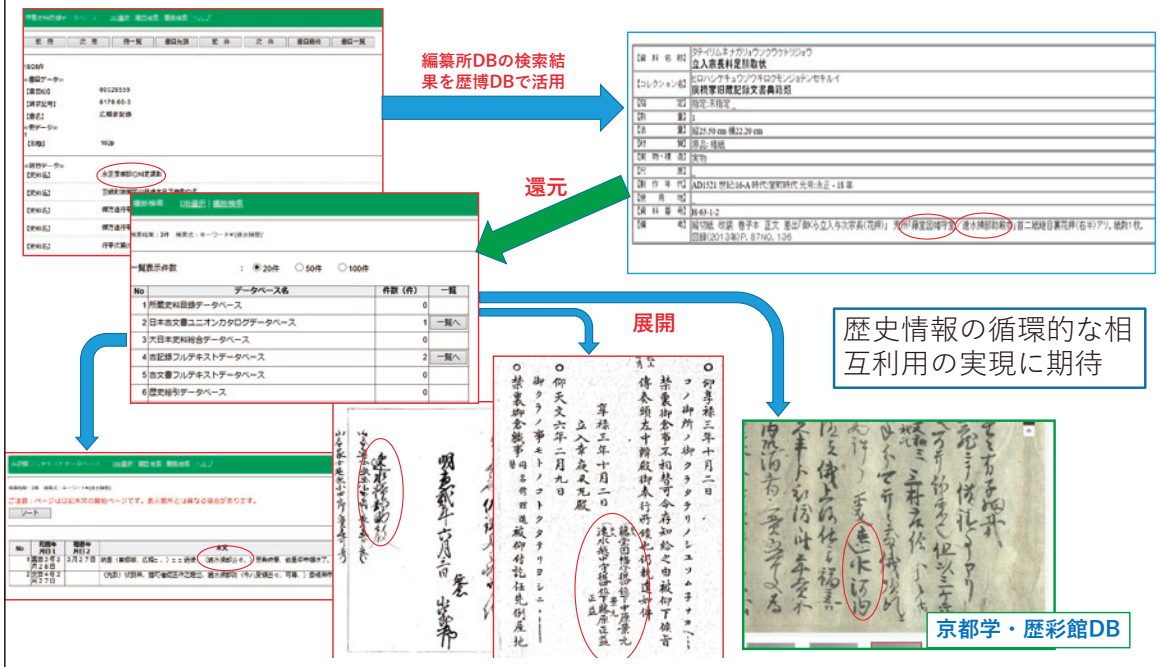
編纂所DBの特性

多様な史料から抽出した膨大な人名・役職名・年代などの情報を蓄積している

所内DBでの横断検索も可能

他機関DBにもリンク

9



10

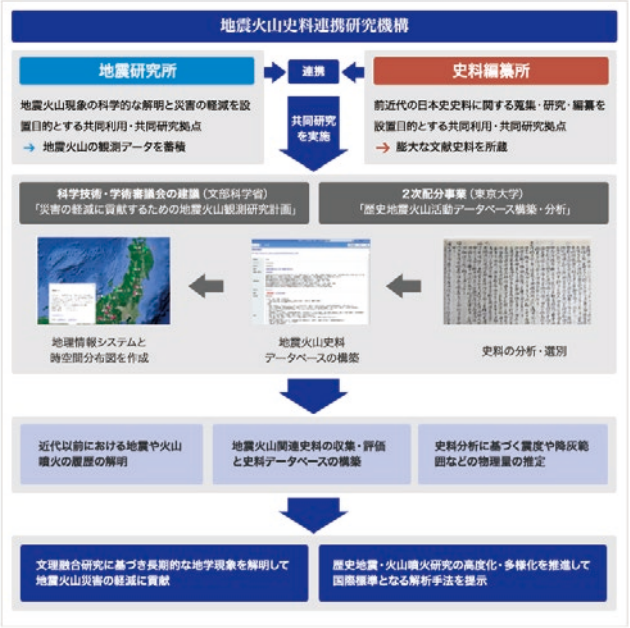
3　地震史料情報収集のための連携

東大地震火山史料連携研究機構

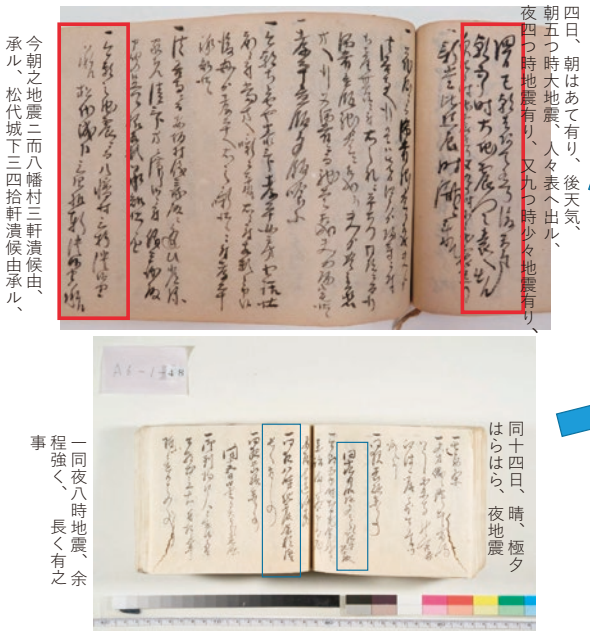
2017年設置
史料編纂所と地震研究所の連携

地震・噴火史料による地震・火山研究

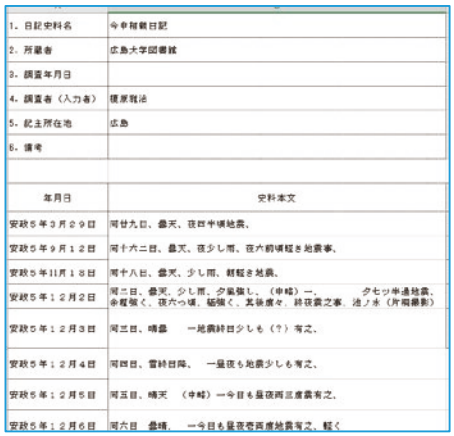
現在は日記から19世紀の地震動の記録を収集
（安政南海トラフ地震の前兆的な現象を探る）



11



調査シートの作成

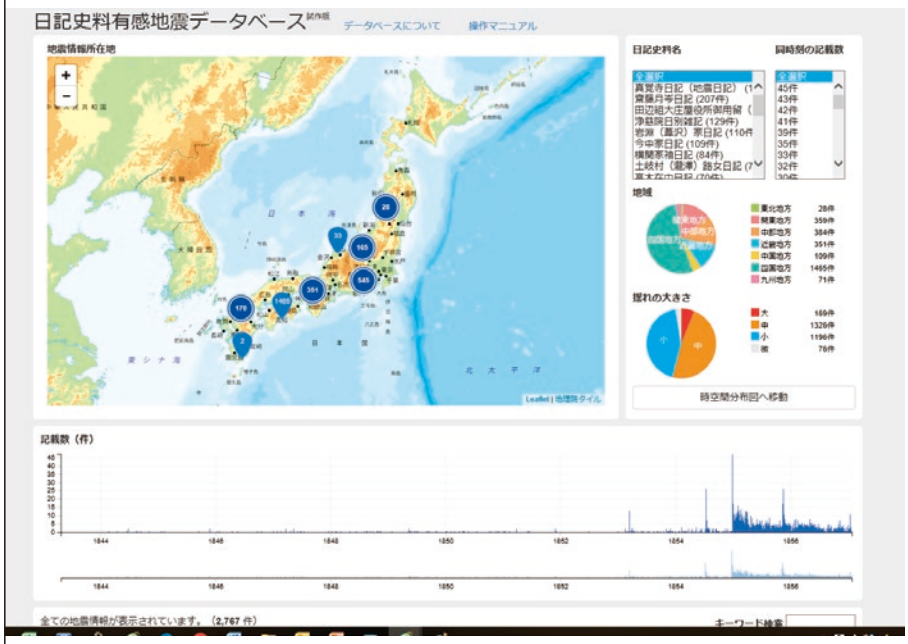


1. 日記史料名	今中陣屋日記
2. 所蔵者	広島大学図書館
3. 調査年月日	
4. 調査者（入力者）	横原親治
5. 記主所在地	広島
6. 備考	

年月日	史料本文
安政５年３月２９日	同廿九日、曇天、夜四半頃地震、
安政５年９月１２日	同十六二日、曇天、夜少し雨、夜六前頃軽き地震事、
安政５年11月１３日	同十八日、曇天、少し雨、軽軽き地震、
安政５年１２月２日	同二日、曇天、少し雨、夕風強し、（申刻）一、　　夕七ッ半過地震、余程強く、夜六ッ頃、強強く、其後度々、終夜震之事、池ノ水（片岡撮影）
安政５年１２月３日	同三日、晴曇　一地震終日少しも（？）有之、
安政５年１２月４日	同四日、雪終日降、　一昼夜も地震少しも有之、
安政５年１２月５日	同五日、晴天　（申刻）一今日も昼夜両三度震有之、
安政５年１２月６日	同六日　曇晴、　一今日も昼夜巻西度地震有之、軽く

DBの構築

12



13

地震史料収集の課題

- 空間的に密度高く収集すること
- 長期にわたる情報を収集すること

- 近世の地方史料の収集は史料編纂所の守備範囲ではない
  各自治体の資料館・教育委員会ごとに把握されている

- 国文学研究資料館、史料保存ネットワーク事業との連携に期待

14

商用による情報提供と公的な情報提供の関係、未来

～ジャパンナレッジと各種連携の考察～

株式会社ネットアドバンス　田　中　政　司

ジャパンナレッジは事典辞書を中心とする商用（有料）のデータベースである。サービス開始から約20年を迎え、大学や研究機関、公共図書館など、世界中で約800を超える機関利用、個人利用のお客様を中心にサービスを提供してきた。また、８年ほど前からは、雑誌記事のアーカイブや歴史的文献など専門性の高いコンテンツも提供し、人文系のデータベースとしては、一定度の評価を得るサービスに成長したと自負している。

そうした活動の中、ここ５-６年、様々なレベルで他機関との連携を行う機会が増えてきた。具体的には、公的研究機関とのシステム間での連携やコンテンツの共同開発・提供がそれにあたる。また、商用のサービスではあるが、世界的に広がる統合的な検索サービス（Discovery Service）などにも積極的にも参加してきた。

本発表では、そうした他機関との連携がジャパンナレッジのサービスにどの程度のインパクトや影響を与えたのかを検証し、今後、商用サービスと公的機関がどのように連携を進めていけるのか、その可能性を探ってみたい

JapanKnowledge

人間文化研究機構　資源共有化研究会

「商用による情報提供と公的な情報提供の関係、未来」

～ジャパンナレッジと各種連携の考察～

2019/2/15

1　1

JapanKnowledge

ジャパンナレッジ（JK）とは？　①

辞書・事典などのレファレンス資料を中心に、学習や研究で使用頻度の高い人文系資料類を検索・閲覧できる有料のデータベース・サービス。2001年にサービスを開始した。法人向けと個人向けのサービスを展開している。

2　2

JapanKnowledge

JKとは？　②

現在、国内外の約800の機関にサービスを提供中。大学や研究機関、図書館のほか、マスコミや企業などで利用されている。海外の日本研究機関での契約が比較的多いのが特徴。

海外での契約機関（契約機関数：110　利用機関数約：130機関）

北米			欧州	アジア	オセアニア
アイオワ大学	デンバー大学	カンザス大学	エディンバラ大学	国立政治大学	オークランド大学
アマースト大学	トロント大学	コーネル大学	オックスフォード大学	国立台湾大学	オーストラリア国立図書館
アリゾナ大学	ニューヨーク大学	コロラド大学ボルダー校	カーディフ大学	輔仁大学	メルボルン大学
アリゾナ州立大学	ノースウェスタン大学	コロンビア大学	ケンブリッジ大学	文藻外語大学	
アルバータ大学	ノースカロライナ州立大学	シカゴ大学	シェフィールド大学	中央研究院	
イエール大学	ノースカロライナ大学チャペルヒル校	ジョージタウン大学	大英図書館	華僑大学	
イリノイ大学	ノートルダム大学	ジョージワシントン大学	マンチェスター大学	広東外語外貿大学	
インディアナ大学	ハーバード大学	ジョンズ・ホプキンス大学	ロンドン大学	香港中文大学	
ウィスコンシン大学	バージニア大学	シンシナティ大学	言語・文化研究図書館(BULAC)	香港大学	
エモリー大学	ハワイ大学	スタンフォード大学	コレージュ・ド・フランス	上海外国語大学	
オクラホマ大学	ヴァンダービルト大学	ダートマス大学	リヨン第3大学	北京師範大学	
オハイオ州立大学	ピッツバーグ大学	テキサス大学	ライデン大学	北京大学	
オレゴン大学	ファーマン大学	デューク大学	ベルリン州立図書館	韓国国立中央博物館	
オークランド大学	フロリダ大学	マサチューセッツ大学	ジュネーブ大学	順天大学校	
UC アーバイン校	ブラウン大学	ミシガン州立大学	チューリッヒ大学	東北亜歴史財団	
UC サンタクルーズ校	ブリガムヤング大学	ミシガン大学	ゲント大学	延世大学校	
UC サンタバーバラ校	ブリティッシュ・コロンビア大学	ミドルベリー大学モントレー国際問題研究所	ルーヴェン・カトリック大学		
UC サンディエゴ校	プリンストン大学	南カリフォルニア大学	ノルディック・アジア研究所		
UC デイビス校	ペンシルベニア州立大学	ミネソタ大学	リュブリャナ大学		
UC バークリー校	ペンシルベニア大学	メリーランド大学			
UC ロサンゼルス校	ボウドイン大学	ユタ大学			
UC マーセド校	ボストン大学	ワシントン大学			
UC リバーサイド校	マギル大学	ワシントン大学セントルイス校			
ミネアポリス美術館	米国議会図書館				

3

3

JapanKnowledge

JKとは？　③

主なコンテンツとしては、レファレンス系の大型辞事典のほか、語学系辞書。専門性の高い「群書類従」「鎌倉遺文」「週刊東洋経済デジタルアーカイブズ」といった資料もオプショナルで追加可能。

日本大百科全書

ランダムハウス英和大辞典
ロベール仏和大辞典

国史大辞典

新編国歌大観と古語大辞典

日本国語大辞典

鎌倉遺文

群書類従

日本地名大辞典（KADOKAWA）と日本歴史地名大系（平凡社）

東洋経済
デジタルアーカイブズ

新編　日本古典文学全集

4

4

JapanKnowledge

JKの特徴（まとめ）

・百科事典を含む辞事典を中心にしたコンテンツのため網羅性が高い。

・利用者は初学者（学部生）から研究者まで幅広い。個人利用では研究職やジャーナリスト、翻訳家などプロフェッショナルユースが大半。

・出版されているものがベースなので、すでにテキスト化された資料が多く、資料の全文検索が可能。

・レファレンス資料のため、専門資料への"入り口"として利用しやすい。

・運営会社であるネットアドバンスはアグリゲーターとして、出版社や機関にこだわることなくコンテンツの調達が可能。

5

5

JapanKnowledge

JKと公的機関のDBの親和性

	JK	公的機関データベース
コンテンツの中身	網羅的、統合的	専門的、独立的
資料としての価値	一般的・汎用的	貴重性高い
利用者	初学者～専門家	専門家
サービスの目的	ビジネス性高い（採算性重視）	公的機関としての使命重視（採算は度外視？）
サービスの継続性	採算が取れるうちは継続	予算がついている間は継続

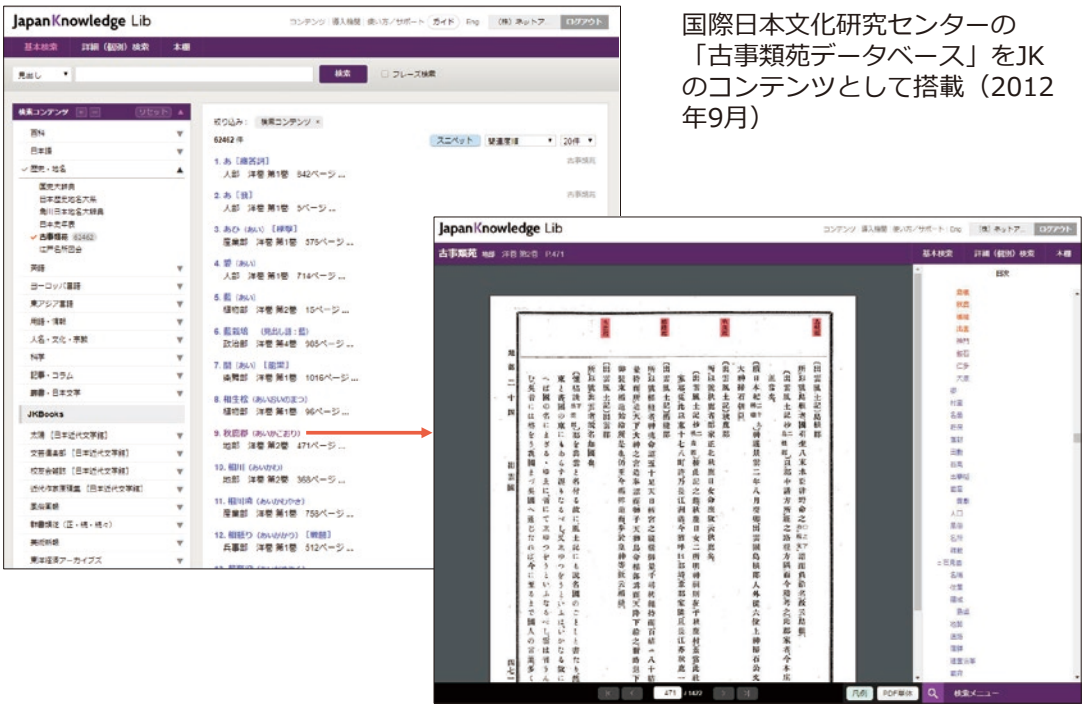
人文系のコンテンツに限っていえば、互いのDBの特徴に補完的な関係が見られるので、連携の親和性が高い。

6

6

JapanKnowledge

具体的な連携①：古事類苑をコンテンツとして搭載



国際日本文化研究センターの「古事類苑データベース」をJKのコンテンツとして搭載（2012年9月）

7　7

JapanKnowledge

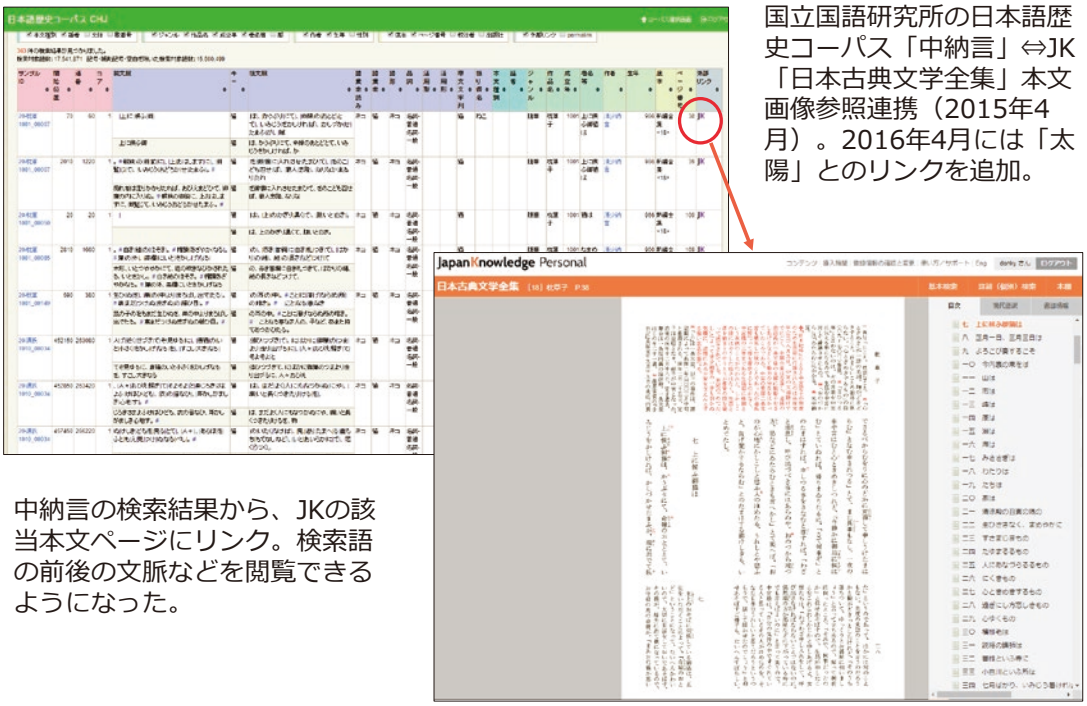
具体的な連携②：NDLサーチ連携

NDLサーチのインデックスデータに、JKの項目データを搭載（2014年4月）

8　8

JapanKnowledge

具体的な連携③：日本語歴史コーパス連携



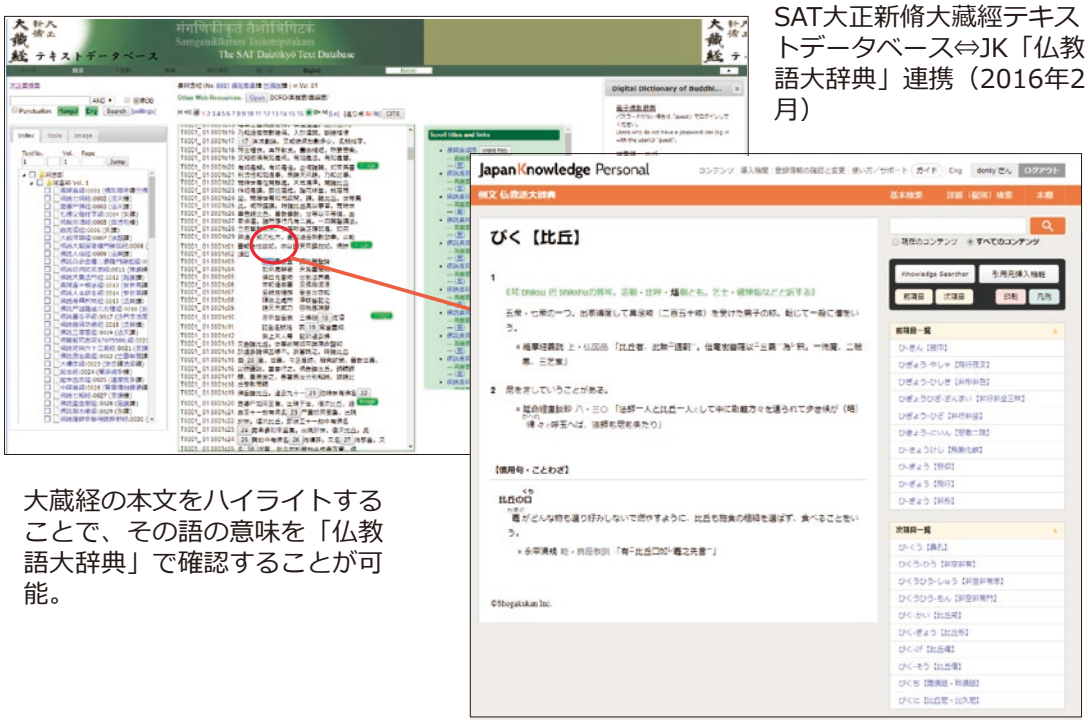
国立国語研究所の日本語歴史コーパス「中納言」⇔JK「日本古典文学全集」本文画像参照連携（2015年4月）。2016年4月には「太陽」とのリンクを追加。

中納言の検索結果から、JKの該当本文ページにリンク。検索語の前後の文脈などを閲覧できるようになった。

9　9

JapanKnowledge

具体的な連携④：SAT大正新脩大藏經テキストデータベース連携



SAT大正新脩大藏經テキストデータベース⇔JK「仏教語大辞典」連携（2016年2月）

大蔵経の本文をハイライトすることで、その語の意味を「仏教語大辞典」で確認することが可能。

10　10

JapanKnowledge

具体的な連携⑤：東京大学史料編纂所連携

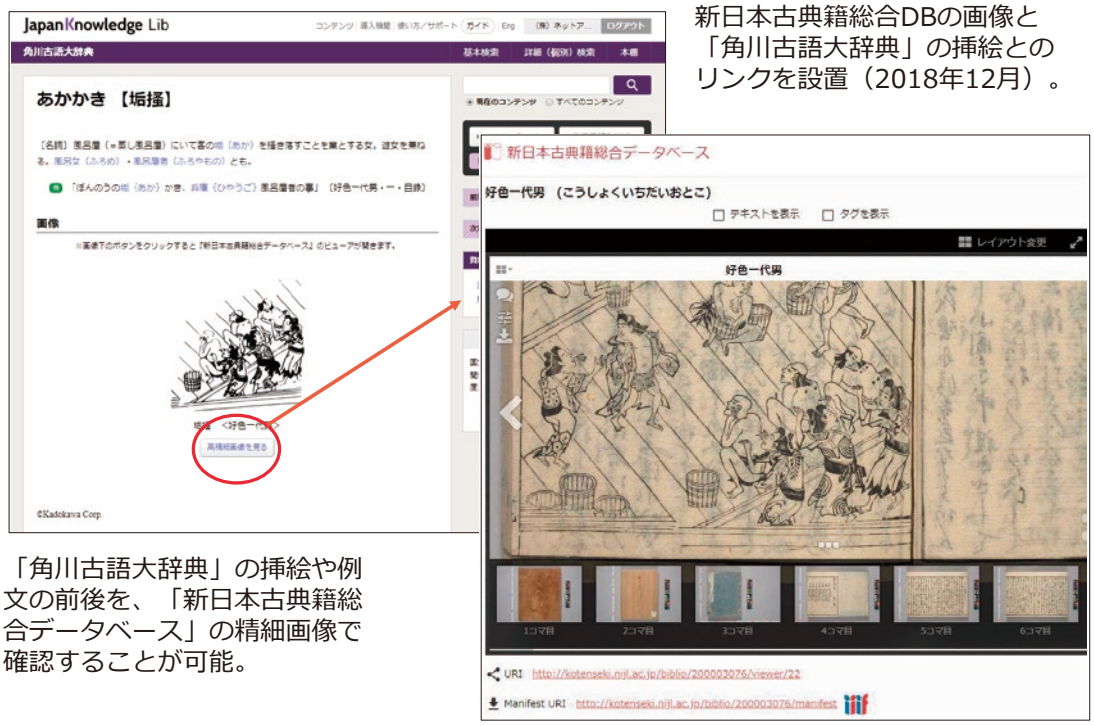
東京大学史料編纂所「日本古文書ユニオンカタログ」⇔JK「鎌倉遺文」連携（2018年12月）。東京堂書店、史料編纂所、鎌倉遺文研究会のコラボレーション企画。

11

11

JapanKnowledge

具体的な連携⑤：新日本古典籍総合データベース連携



新日本古典籍総合DBの画像と「角川古語大辞典」の挿絵とのリンクを設置（2018年12月）。

「角川古語大辞典」の挿絵や例文の前後を、「新日本古典籍総合データベース」の精細画像で確認することが可能。

12

12

JapanKnowledge

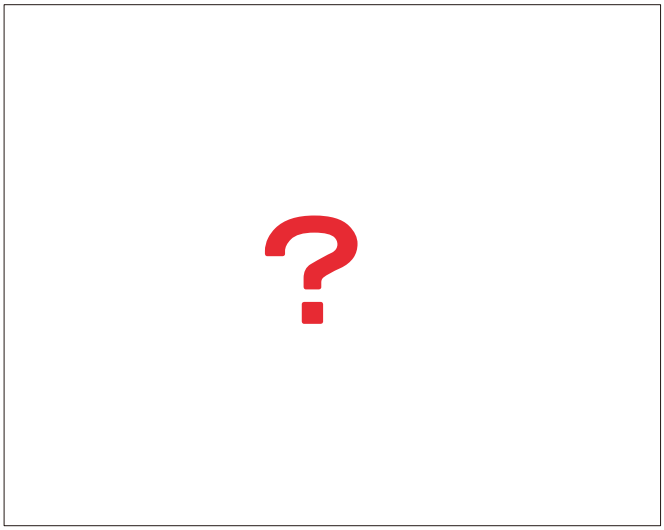
具体的な連携⑥：新日本古典籍総合データベース連携

「新日本古典籍総合データベース」のインデックスをJKの検索インデックスに統合。
JK内では書誌情報の一部を確認できる。

13 13

JapanKnowledge

具体的な連携⑦：歴史民俗博物館「Khirin」



歴史民俗博物館「Khirin」⇔JK「日本歴史地名大系」「角川日本地名大辞典」連携（詳細検討中）

14 14

JapanKnowledge

公的機関との連携のタイプと目的

■本文ページ同士のリンク
⇒　利用者の利便性向上、相互の利用率向上
（国研「日本語歴史コーパス」、国文研「新日本古典籍総合データベース」）

■検索APIの提供
⇒　利用者の利便性向上、JKの利用率向上
（国文研、一部の大学）

■インデックスデータの提供・共有化
⇒　JKへの流入を増加
（ SAT大正新脩大藏經テキストDB、NDLサーチ連携）

■コンテンツをJKに提供いただく
⇒　ロイヤリティや使用料をお支払いする
（日文研「古事類苑」、史料編纂所「鎌倉遺文」）

15　15

JapanKnowledge

JKにおける各連携の実績③

結論として、我々のDBにおいて、

★「古事類苑」搭載

★ 日本語コーパス（中納言）連携

の二つが、安定的に一定の成果挙げていると考えている。

↓

「古事類苑」については、コンテンツ自体の価値と、JKの他コンテンツとの一括検索の利便性が、利用が伸びている原因と考えている。

「日本語歴史コーパス」連携については、コーパスで調査した内容を、全体の作品で俯瞰する必要があるなど、お互いのDBでは完結できない機能を補完し合い、相性が良い。

16　16

JapanKnowledge

展望（希望？）と課題①

・我々商用DBにとって、公的機関との連携は単に利用率を向上させるためだけの施策ではない。最近は仕様が変更されたようだが、政府機関や大学との連携は、GoogleやYahoo!の検索順位を上げるという実利もあった。実際に海外の大学では、こうした研究機関との連携を高く評価し、契約に至ったケースもある。

・ネット上の辞事典は「簡便なもの＝無料」「専門性の高いもの＝有料」の棲み分けができつつある。有料DBである我々のビジネスを考えると、専門性が高いコンテンツを増やすことがDBの価値を高めていくことであり、外部の専門的DBと機能的に補完し合える連携はその価値をさらに高めることにつながると考えている。

・DBは単に資料を利用するという用途から、「テキストデータの活用」「検索結果の分析」といった需要が出始めている。そのような中で、単なる本文ページの相互リンクといった連携では利用者は満足しなくなっていくのではと危惧している。より付加価値が高い機能連携を考えていく必要があり、そのためにも、官民の枠を超えて情報共有ができる場の必要性を感じる。

17 17

JapanKnowledge

展望（希望？）と課題②

・商用DBと公的機関のDBでは、対象とする利用者や目的が違うため、データ形式であったり、認証の仕組みが自ずと異なる。そうした差異を利用者に意識させることなく開発を進める必要があるが、いまだハードルが高い。会員認証の仕組みを意識させないことや、オープンソースを前提としたIIIFなど、業界標準になりつつある規格を、我々のサービスにもどう取り入れていくか、今後の連携に考えるとまだまだ課題は多い。

・開発の段階から、他サービスとの連携や、データの二次利用を見越した著作権処理が必要。我々のサービスでも起こりがちだが、連携先にデータを渡す際に、著作権の制限で必要な情報を利用できないということがあり連携に支障をきたしたことがあった。

18 18

ディスカッション
「人文学データの今後と人間文化研究機構が目指す情報基盤」

司　　会：関　野　　樹（国際日本文化研究センター）
パネラー：永　村　　眞（人間文化研究機構）
大　内　英　範（人間文化研究機構）
菊　池　信　彦（関西大学アジア・オープン・リサーチセンター）
榎　原　雅　治（東京大学史料編纂所）
田　中　政　司（株式会社ネットアドバンス）

（関野）　司会を務めます国際日本文化研究センターの関野でございます。よろしくお願いいたします。

発表の中で質問の時間が取れなかったので、ディスカッションを始める前に、少々質問のお時間を最初に取りたいと思います。どの発表にでも結構ですので、ご質問もしくはコメントがありましたら、お願いします。いかがでしょうか。発言される際は、お手数ですが所属とお名前をおっしゃってください。

（永崎）　人文情報学研究所の永崎と申します。今日は大変いろいろ勉強させていただきまして、ありがとうございます。ようやくこういうところまで来たというのだなと改めて感動しているところでして、関係者の皆さまのご尽力に改めて感謝しております。

ここまでようやく便利に使えるように、サクサク検索できるようになってきたところで、特に大内先生のお話、それから菊池先生のご指摘も踏まえた上で、私の方でもう少し、もしかしたら検討された結果、今があるのかもしれないですが、もし検討されていなければぜひお願いしたいと思いましたのが、大学図書館とどういう関係にあるのかというところです。

大学図書館の方では、JPCOARというのはご存じですか（少し間を置き、大内先生がうなずいた後）、大内先生は。JPCOARの方で、例えば研究データを集めてデータリポジトリとして、データをどんどん研究者から集めようとしているのですけれども、実際にそこにはまだそんなにきちんとデータが集まっているわけではなく、掛け声がどれぐらいちゃんと広がるのかという面はあるにせよ、本来、nihuINTあるいは機構の各機関がやっている動きと結構近い話が多い。というのは、nihuINTに集まっているデータベースというのは、恐らく研究成果的なデータベースが多いのではないかと思うのです。ジャパンサーチで検索してもらえると、それはそれでいいと思うのですけれども、オープンサイエンスという文脈でつながっていくことで、またちょっと違う文脈の話が出せるのかなという気がしておりまして、そこが、特に大学図書館業界というのは研究者業界と違って、組織として職掌を明確にしてこういう仕事をしてくださいと言うと、それに対するものを出してくれるので、研究者のグループの仕事とは、ちょっと深さがどうかという問題が常にあるのですけれども、そこら辺がちょっと違うところがあるかと思っています。

そこのところをどういうふうに、何か見通しがおありか、それとも、やったけれどもあまりうまくいかないからもうやらないのか、それとも今後可能性があるのかというあたりを。JPCOARのあたりとうまく連携できれば、個々の教員とnihuINTの関係がつくれる。それで、nihuINTのデータをハブとして各大学の教員にまでつながっていけるというルートができるのではないかというところが期待されるところなのですが、いかがでしょうか。

（大内） ありがとうございます。図書館関係の研究会にはなるべく参加するようにしているのですけれども、まだそのような、私たちと一緒にやりましょうというような話は検討したことがないのではないかと思います。せっかくそういうご示唆を頂きましたので、もし何か話ができるのであれば、させていただきたいと思います。

（永崎） もし、まだちゃんとしていないのであれば、彼らも既にメタデータスキームなどを作ったりしているので、そこら辺のところからきちんと入れると思うのですね。そうすると、彼らは結局、大学の中で教員からの支持を得ていない動きになってしまっているので、しかし、これが、nihuINTから人文系のメタデータに関してこういうふうに貢献しているという話ができれば、向こうも随分また違うアピールの仕方ができるようになってくるかもしれなくて、そうすると、nihuINTと大学教員が自分のところのリポジトリに入ったデータの連携とかがすごく深くできるようになる可能性もあるので、それはぜひご検討ください。というか、ぜひ前に進めてみていただけるといいのではないかと思いますので、よろしくお願いします。

（大内） ありがとうございます。

（永崎） 今日はその話をした方がいいのかもしれないと思って。もし検討していなかったら、ぜひ、しなければと思って来ましたので。はい。ありがとうございます。

（大内） はい、ありがとうございます。まずは意見交換というか、情報交換というようなところだけでもやってみたいと思いますが、菊池先生から、何かこの点についてご意見をうかがえませんか。

（菊池） 永崎先生のおっしゃるとおりだなと思って、今、うなずきながら聞いていたのですけれども、ちょっと前にSPARC Japanで人文系のデータ、研究データの扱いをどうするのかというイベントがありまして、そこの議論の中でも、やはり「人文系のデータがなかなか出てこないよね」というところを皆さん議論していらっしゃったので、大学図書館関係とか、そういう方面の方々が。その文脈でもやはりJPCOARの話が出ていたので、nihuINTがその人文系のデータを扱うハブ的な立ち位置に立つのであれば、やはりそことの連携は、積極的に議論を進めていただくのが結構いいのかなと思いました。すみません、乗っかるだけでしたけれども。以上です。

（関野） ありがとうございました。他のお二方は、この点に関して何かコメントなどはありますか。フロアからはいかがですか。よろしいでしょうか。

それでは、他に何か質問、コメントなどがありましたらお願いいたします。人数も少ないので気楽な形でディスカッションを進めていきたいと思います。フロアの方もどんどん積極的に参加していただければと思います。

それでは、ディスカッションに移りたいと思います。いろいろ興味深い話、参考になる話を伺って、私としても勉強になったところです。最初にお伺いしていきたいところは、いろいろな形でデータベースやアーカイブを公開している現状を紹介していただきましたけれども、利用者のニーズをどのように皆さんは捉えているのか、もしくは、くみ取っているのかというところがわれわれにとっても今後どのように開発を進めるかという点で大きな問題になってくるかと思います。大内さんからの説明にあったように、例えばnihuINTが現状のユーザーインターフェースを縮小していて、APIに置き換えてゆくということになると、ユーザーのニーズを捉えることは諦めてしまって、実際にユーザーが触る部分はユーザーに丸投げしているというふうにも捉えられかねない。ある程度はそれでできるところはできるでしょうし、できないところもやはり出てくると思います。ユーザーのニーズという点で、ユーザーはデータをどのように使うのか、例えば、データがどのようにリンクされると、ユーザーは使いやすいのかといったニーズは、どういうふうにくみ取ってこられたのか。そして、それがどのような形でご紹介いただいたような形になったかということについて、もう少し具体的な事例などをお話しいただけると、こちらとしても参考になります。いかがでしょうか。

発表順でよろしいですか。では、菊池先生から。

（菊池） うちが作るデータベースとして、ユーザーは、まず教員だと思います。教員の研究からスタートしていくので、教員の研究に関心のあるようなユーザーが外側に広がっている。図書館関係ですとか、あるいは一般のユーザーという形になるでしょうね。ニーズをどう組み込むかというのは難しいなと思います。特に、中国の研究などのテーマになっていますので、日本に対して訴えかけるものがどれだけあるのかなというのは未知数。というか、むしろニーズのないところにニーズを喚起させてデータを提供していくような姿勢が必要なのかなと思っています。

（関野） 菊池先生、ありがとうございます。

（榎原） 史料編纂所のデータベースのユーザーの中心は研究者だろうと思いますが、アクセス数は月にアクセス30万件と出ていますので、一般の方も多分アクセスしているのだろうと思います。ただ、編纂所で公開しているデータは一般向けのデータには決してなっていないです。研究者でないと十分使いこなせないような内容になっていると思います。歴史的な一般的なことを知りたいと思って編纂所のデータベースを引く方はいると思いますが、多分その要請には応えられていないだろうと思います。最近、いろいろな大学のデータベースを見ると、一般的な質問に答えるようなコーナーが設けられているようですが、そういったニーズに応える用意は、編纂所のデータベースはなっていない

と思います。そのあたりは、これからの課題かと思います。

（田中） われわれは商用データベースですので非常にシンプルです。お客さまからTwitterやメールなど、いろいろな形で要望が上がってきて、こういうコンテンツを増やしてくれとか、ここが使いづらいだとか、ここをこうしてくれだとか、さまざまな意見をいただきます。そういう声を斟酌（しんしゃく）しつつ、かけたコストに見合ったリターンがあるのかを考えながら、「これはいけるね」ということであればコンテンツを追加したり、連携をしたり、機能改善したりとかということを日常的に行っています。ですので、われわれは他のデータベースと比べて、お客様と非常にシンプルな関係性を維持しながら作っているのかなと感じました。

（関野） もう少しこの話題を深掘りしてみたいと思います。得てして研究者がデータベースを作ると、その研究者にとって使いやすいデータベースだったり、データの構造だったりになりがちですよね。それは僕自身もよく経験しますし、皆さまも思い当たるところがあるのかもしれません。一方で、大内さんからも少し説明がありましたけれども、思いがけない発見をするといったときには、結局データが発見する対象につながっていかないわけですよね。データだけ丸投げしてしまって、あとはユーザーが使いやすいようにというのも一つの方法かもしれませんが、データをつなげるという点で、何らかの工夫やお考えはありますでしょうか。元のデータに対して、もう少し違う見方ができるような考え方やアイデアがあれば、実現していなくても結構ですので、お聞かせください。

（榎原） 意図せずして思いがけない発見ができるというのは、あらゆるデータベースについていえることで、とりわけ工夫をしなくても、ある程度、実現できているのではないかと思います。今日の報告を用意するために検索をちょこちょことやっているだけで、あれ？と思うような発見があり、結構面白かったです。たとえば歴博のデータベースを引いてみたら、備考欄とか参照欄にいろいろ書いてあることが目に入ってくる。それだけで十分意外な発見ができるのですね。いろいろな情報がたくさん入っていれば、それだけでそういうことは果たせるのではないかと思います。

（関野） 菊池先生は何かありますか。

（菊池） 言われるがままに作っているところがあるのですけれども。考え方としては二つあって、一つは、思いがけない発見をというスタンスで情報学的な見地からてこ入れをするということもあると思うのです。それはそれで大事なのですけれども、一方で、研究者が自分のために作るというのは、作品としてのものということもあると思うのです。特に歴史研究者だと、論文なり本なりを作るのと同じようにデータベースを作る。歴史叙述としてのデータベースというのも、私はあり得ると思っているのです。ですから、そのスタンスに立てば、研究者が自分の考えでデータベースを作っていくというのも、決して否定されるようなものではないのではないかなと思っています。

（関野） ありがとうございます。田中さんの話を聞いていても思ったのですけれども、ジャパンナレッジでは、使われることを前提とした取り組みやデータベースの公開が見ていてよく分かりました。一方で、菊池先生のおっしゃるとおり、やはり研究者というのはデータベースを作ること自体も一つの目的としていて、それを公開することが論文を書くのと同じような位置づけになることがある。そこは商用ベースのデータベースと少し違います。ジャパンナレッジでは、データベースについて研究者からどうしてもこうしてほしいとか、ああしてほしいという要望があったりする場合に、何かされているのでしょうか。

（田中） 先生方からご要望があるもので、今、非常に多いのは、テキストのダウンロードを可能にしてくれというものです。これは、版元である出版社と交渉して許諾を取るのですが、やはりそこは出版社の命でもありますので、なかなか実現できない。ただ、われわれとしては一応、何らかの形で研究者の方、学生さんも含めてですけれども、新たな研究の要望に応えていかないといけないだろうということで、テキストデータをお客さんに渡さない形でテキストマイニングができるような仕組みというのを、今、研究はしているのですけれども、まだ発表はできていないという段階です。

（関野） 大内さんの方からは、新しいnihuINTについて、いくつかアイデアをご披露いただきましたけれども、この点については何かありますか。

（大内） データベースがたくさん存在していると、データベースごとの特徴を知らないと、なかなか求める情報にたどり着けないということが結構あって、これを調べたいなと思ったときに、初めてのデータベースを使って、目的の情報にたどり着くというのは、やはり幸運が必要なことが多いとは思うのです。そういう意味でいうと、nihuINTには機構の中のデータベースだけでも150ぐらいのデータベースがDublin Coreを基礎とするメタデータで整理された状態で入っているわけですけれども、むしろ先ほどの報告の中でも言いましたように、分野ごとにユーザーインターフェースがあってもいいのではないかということです。それはこちらでは手当てできそうもないので、それなりの人たちがたくさん各分野にいるだろうから、そういう人たちに作ってもらったらというような話なのですけれども。各分野の方々が必要としているだろうというのを見越した上でユーザーインターフェースを作ってもらって、なるべくその分野の人がたどり着きやすいような、求める人がたどり着きやすいようなものを作っていただくことで、ある程度、求める情報にたどり着くまでのハードルも下がっていくところがあるのかなと考えているところです。

また、ニーズの話があったのですけれども、nihuINTもニーズの調査とかをした方がいいのかなと、思ったりしています。なかなか難しいとは思いますけど。これまでそういうのはしたことないですよね。

（関野） 過去にやったことはあったと思うのですが。

（大内） そうですか。本当はそういうのをやって、どういうところにニーズが多いかを把握した上で改善していくのが必要なのかなと、担当としては思っているところです。

（関野） ありがとうございました。フロアの中にも、恐らく自分でデータベースを作って公開されている方は多いと思います。例えばユーザーのニーズをどのようにくみ取っているかとか、逆に、自分のデータベースはこう使ってもらうべきだといった話はあるでしょうか。

（永崎） ちょっとだけ、私のところでやっている話をご紹介させていただきますと、大藏經テキストデータベースというものを作っているのですが、それで大体月間30万件ぐらいのアクセスがあり、世界中の研究者から使われているようなものです。１億字のテキストデータベースの全文検索プラスいろいろな機能ということなのですけれども、世界各地に講習会をやりに行きます。大体、３人以上希望したらそこに行って、私か誰かが使い方を説明する。一通り説明し終わってから、皆さんのご要望をお伺いするということを世界各地でずっと行ってきています。

アンケート調査をやったことがありません。なぜアンケート調査をやらないかというと、使い方をよく知らないでアンケートに答えるというか、アンケート以前の話だったのですけれども、すでに用意されている機能を知らないまま「あの機能がないから付けてくれ」というクレームを以前から結構頂いていて、この状況でアンケートをやったら、使い方を知らないでアンケートに答えて、数字だけ異様に低い何か変なものが出てきてしまうのではないかと。あるいは、使い方を知らないということは、そもそも、その便利さを全く想像もできないわけですから、その便利さを知った上で、さらにこういう要望が欲しい、こういうニーズがあります、というのと、そうではない何も知らない状態でこうだというのが、アンケート調査をすると、一つの平面で同じように語られてしまう危険があります。ということで、とにかく講習会後にアンケートを行うという方式でやっております。

（関野） 他はいかがですか。

（相田） 国文学研究資料館の相田です。昨日、会議がありまして、国文研でやっている新日本古典データベースや画像データベースの国際化をやろうとして、では、それをやろうとするときの「そもそもの国際化は何よ」という話になって、ただ単に英語圏とか、あるいは中国語圏のナビゲーションウインドーを付けるということは既に実験されていたのですけれども、なかなか成果は出ないし、データをローマナイズするということをやってしまうと、「こんなものはGoogleでやればすぐに出てくる」と。ということで、もう一つの観点として、やはり誰でも使えるようなインターフェースをということになってきますと、これは自分の主観なのですけれども、究極は、やはりユニバーサルデザインを目指して、外国人というよりも、言葉の分からない子どもでも引けるというインターフェースをやらないと、そこにアクセスできない。

特に、うちのところで出している画像などをどうやって管理するか。外国人が触れる

イコール子どもでも触れるという、ほぼ同じ意味になるかと思うのです。そういうユーザーインターフェースを作っていくことで、逆に文字のことを考えると、ネットアドバンスさんの方は海外の130機関からのアクセスがあるという。やはり日本語だけでもやっているのかということですよね。中国圏だと、日本語で入れたつもりが実はUnicodeで文字化けするものが結構あるとか。また、ジャパンナレッジは結構外字を使っていますよね。自分自身も個人ユーザーでお金を払ってはいるのですけれども、ただ、今回紹介されているのは個人ユーザーでは手の届かない「ライブラリー」のものばかりが出ていて、非常に悔しい思いをしながら拝聴させていただいているのですけれども。

それはさておき、やはり、いちいち探さなければ、ものによってマニュアルが違うということについてはユーザーにとっては敷居が高くなってしまうと思います。自分も立場上関わっている中で言うのも何ですけれども、nihuINTも、とっかかりのユーザー画面では、勝手を知らないと引き方が分からないというところもあるわけです。大抵は慣れ過ぎてしまって、逆にマニュアルとかを読まなければ触れないというインターフェースを持つものだと、逆にそれが障害になって使われなくなるということもあるかと思うのですが、そのところで何かいいお考えがありましたら。お三方、関西大学という立場からでも聞かせていただきたいのですけれども。

（田中） 基本的にジャパンナレッジは、「ジャパンナレッジ」というぐらいで、日本研究を行っているお客さまが対象です。インターフェースに関しては英語のものを用意はしているのですが、いかんせん、コンテンツの中身が全て日本語ですので、日本語が読めるということ、それから日本語を表示できるシステムが入っているということを前提にサービスを提供しています。ですので、中国語の話しが出てきた時、ちょっと今、耳が痛いなと思いながらお話を聞いておりました。多言語の対応については今後の課題だと認識しています。コンテンツのデータも完全にUnicode化されているわけではなくて、JISをベースにしたデータが多いものですから、中国系の文字の問題というのは、認識はしているのですが、コンテンツも膨大なため、すべてをUnicodeに換えるというのは一朝一夕にはいかず、今後の課題として考えさせていただきたいと思います。

（榎原） 同様ですね。最近仕事上で必要があって、韓国の歴史史料のデータベースをよく見ています。本文は漢字で出てくるので読めるのですが、タグなどは全部ハングルですから、ほとんど読めないわけです。そのあたりを国際的に分かりやすい表現にしてくれればいいのにと思うのですが、振り返ってみれば、自分たちのところのデータベースも同じではないかと思うわけです。データ本文が日本語なのは仕方がないのですが、操作説明であるとか、タグであるとかは英語表記を付けておくべきではないか、よその国のデータベースを使って気づいたところです。

（菊池） 関大の方は、今、まだ作っているところですので、中国語のインターフェースを用意すればいいのかなとか、素朴なところで考えていたのですね。むしろ国文研さんのデータベース、デジタルアーカイブが国際化を目指している、その背景の方が私は気になっているのですけれども、教えていただけないでしょうか。

（相田） これは、予算が付いたときにミッションとして入ってしまっているので。これはぶっちゃけた話なのですけれども。ただ、それをやっているときに、やはり生半可なことをやっていると、使えないものが、現状でもいろいろ実験的に試みられた中でもやはり使えないということがよく分かったということがありましたね。ちょっと視点を変えて、もう少し幅広い層の方から使ってもらえるようなインターフェースというのをどうすればということが今後の課題になるのかなと。

これまでのデータベースは大抵、パソコンの画面でないと探せないものが多かったのですけれども、スマホなどのインターフェースが主流になってきますと、今度は共通するアイコンで探せるようになるか、見れば分かる、あるいは障害者のある人でも触れば分かるというふうなところが実現されるかもしれません。前任者は、くずし字を触覚でちゃんと認識していたとおっしゃって、要するに、塙保己一というのは全盲でしたけれども、くずし字も手で触って読めたことを実証したというようなことを言ってました。そこまでを考えてということになりますと、やはり情報を集めるのに、鈴木先生のところみたいな博物館系とか、あるいは山本先生のところみたいなものがどういうアプローチしているのかということも考えないと、難しいのかなという感覚を持っています。

（永崎） 人文情報学研究所の永崎です。今の話で、ちょっと後で詳しくお伺いしたいのは、今の新日本古典籍総合データベースそのものをアンケートか何かした結果、国際化がまだ不十分だったという、その話は後で詳しくお伺いしたいのですが、今の相田先生のお話は恐らく、ここでやろうとしているのは、研究データのデータベースの連携の話ですので、そうすると、研究方法論がきちんと国際化されて共有化されて、アイコン化できるレベルに抽象化できないと、多分うまくいかないのではないかというか、まずそちらが先なのではないかという感じがします。データベースを作る人が集まって話し合うことではなくて、国文研の先生方が各国で集まっていただいて、英語ではこういう言語にと。こういう日本語での、例えば今、落合先生が書誌学の国際化のプロジェクトをやっておられますよね。ああいうのも、書誌学だけではなくて、国文学全般にわたってきちんと、この用語はこの言語だとこれに当たりますねと。やはり言語によってズレがあるのですよね。今、私はTEIコンソーシアムで国際化の話をやっているのですけれども、そちらはその話は英語圏発の話になっているのですけれども、それもやはり日本語とずれますし、ドイツ語圏ともずれるらしいのですね。ですから、欧米でも別にピタッと合っているわけではなくて。書誌学用語ですら。しかし、それをそろえないと、TEIコンソーシアムはそういう仕事をずっと30年ぐらいやっていますので、そろえないと話にならないので、何とかやっているわけですが、それを日本語にどうやって落とし込むか、まさに国文研のお仕事かと思いますので、ぜひ頑張っていただいて、それをアイコンに落とし込むところまで頑張ってください。

（関野） 永崎先生より非常に有用なコメントを頂いたと思います。確かに、研究者がデータやデータベースを作るのにどのように関わるのかは大きな問題の一つだと考えています。つまり、ここまではデータを使う側としての研究者の話をしてきましたけれども、データを作る側としての研究者の役割も、やはり、重要になってくると思います。今、

永崎先生が話されたように、日本語を英語にするという話だけでも、研究者がきちんと話をしないとデータづくりが始まらないです。そもそもデーベースそのものというのは研究者が持つ知識を固めたものという捉え方もできますから、榎原先生から話があった、いろいろなデータをリンクするという話も、どこに何があるのかという研究者の知識があって初めて成り立つ話なのだと思います。そういったところで、作る側としての研究者がどうあるべきなのか、どういうふうに関わるべきなのかということについて、話を伺っていきたいと思います。

それぞれの機関では、プロジェクトや個々の研究者の研究成果としてデータベースが作られていることと思います。では、具体的にどのような経緯や過程でそれらデータベースが生まれてくるのか。また、データベースができてくる中で、研究者にどのような関わり方をするのが今後のデータベースを作っていく上で望ましいのかということについて、少しご意見を伺っていきたいと思います。それでは、田中さんからお伺いしてよろしいですか。

（田中） われわれのデータベースの元となる資料は、どちらかというと本になる段階で、研究者の方が関わられますので、われわれが実際にデータベースを作る段階では、それほど研究者の意見は反映されません。ただ、今回作成した「Web版鎌倉遺文」では、元データをお持ちの史料編纂所の先生方の意見をかなり取り入れるなど、われわれの方で枠組みの提案はするのですけれども、確認作業を一つ一つ入れながら、アドバイスを頂きながら進めています。

（榎原） 歴史の分野でデータベースというものが話題になった最初のころは、いろいろな史料の索引を入れましょうというところから始まったと思います。これは、今日いらっしゃる永村さんが先導されたわけですけれども、そのうち全文テキストをどんどん入れていくようになりました。単独のデータベースの形式としては、すでに完成に近い水準に達しているのだろうと思います。これからは、どういうデータベースを作るかというのを越えて、どうつなげていくかというところを具体的に構想していくことが必要なのではないかと思います。それを考えるためには、研究者自身が実際にいろいろなデータベースを使いこなしていく経験を蓄積しないといけないだろうと思います。

（菊池） うちでできているかどうか、これから本当にできるのかどうかはさておき、先ほど歴史叙述としてのデータベースという言い方をしましたが、データを使ってどう表現するか。可視化とかビジュアライズとか、いろいろな表現がされていますけれども、データを使って歴史を書くという行為をどう表現していくのかというのが確かに必要なのかなと。整理をしていくとか、検索をしていくという、研究成果として出す前の前段階のところをやっていくだけではなくて、まとまったものをどう表現していくのか、分析した結果どう表現して、今まで書いていたこと、本や論文で書いていたのではない別の表現が歴史家としてできるのではないのかというところに、これから向かっていくのではないかと思います。

（関野） ありがとうございます。データを作る上では、それぞれの研究者の貢献もあります。一方で、大内さんからも話がありましたが、人間文化研究機構で取り組んでいる地名辞書のようなデータを作るためのデータというのも必要になってくるのだと思います。そういった基盤データは一体誰がどのように作っていくのかも一つ課題です。これは、人間文化研究機構の役割なのかもしれません。

この点について、こんなデータがあったらこんなことができる、こういうデータがあれば新たな研究が展開する、もしくは、榎原先生が挙げられた地名のデータのように、こういうデータがあれば、データ同士を連携できるというような、基盤的なデータとして、思い当たるようなものはありますでしょうか。もしありましたら、フロアも含めてどなたでも結構ですが、いかがですか。恐らく、こういった基盤データは、データベースというよりは、菊池先生がおっしゃったようなインフラ的な、図書館に近いような機能になるのかもしれません。論文を書くために使う専門のデータベースというよりは、どちらかと言えばインフラ的なものになるのかもしれません。永崎先生、どうぞ。

（永崎） すごく手短に。そういう意味では、やはり日本の全部の古典籍のテキストデータがばーっとあって、そこにいろいろな資料を生やしていけるような感じになるといいなと思っています。

（関野） 何か、元になるプラットフォームになるようなデータベースのデータがあって、そこに枝葉が伸びていくようなイメージですね。そうすると、幹に相当するものは一つだけではなくて、幾つもの木が立っていて、それぞれに枝葉が伸びて、それらがまた絡み合ってというようなイメージになるのでしょうかね。

（永崎） そうですね。テキストというと、特に古典テキストはどうしても本文を定めるは難しいですけれども、一つテキストのデータがあって、そこから、どれが本文か分からないけれども、この写本ではこうなって、この写本ではこうなってというのが、さらにそこから、あちらにはこう引用され、こう引用され、こう発展してとか、何かそういうのがばーっと日本古典籍であるとか、うちの仏典のデータベースでやっていって、仏典のデータベースの方はそういうふうになっていて、それでこの世界中の資料がつながっている状態なのですね。ただ、それでわれわれは一生懸命やっているのですけれども、例えば日本霊異記とかに始まり、いろいろなところに本当はつながるはずなのですけれども、つないだら、それはそのまま世界のチベット語のデータベースも、サンスクリット語のデータベースもばーっとつながるはずなのですけれども、ただ、そこが自由にうまくつなげられるところがなくて「うーん」というところもありまして、そういう安定したテキストデータの場があるといいなと思います。

（関野） 他はいかがですか。何かこういうデータさえあれば、もっとこういうデータが作れるのにとか、何か思いがあるのかなという気が僕はしてならないのです。もちろん、菊池先生も榎原先生も大内さんも、一研究者としての立場でも結構です。これがあればもっと自分の研究が進むのにというデータがそれぞれあるのではないかと思うのですけ

れども、すぐには出てこないでしょうか。

（榎原） 個別の文献に即して、あのテキストデータが欲しい、これが欲しいと言い出せば、これはもうきりがないが、そこはあえて申し上げない。全文テキストを供給する仕組みができているというところで、今のところいいのかなと思います。地名については、文献テキストに地名を自動的に組み込むシステムができればとてもありがたいですね。地名に位置情報を付与することは、関野先生たちのご努力でできたわけですけれども、それを文献テキストの中の地名に自動的に組み込んでいくことはまだできないので、そこができるような仕組みができればとても有用だと思います。

（相田） 国文学研究資料館で最も多く出ているのは、書籍のデータです。これは古典籍、いわゆる『国書総目録』から起こったものがあるわけですけれども、それでビブリオ ID とかデータ ID が付されてあって、異表記のデータを管理することは大本のところへたどっていって、そこで管理される ID ナンバーからデータをたどっていっている。問題は、かつて歴史人物データをやったときの経験から言うのですが、やはり表記、特に歴史的な人物になってくると、秀吉のように姓がコロコロ変わって、名も変わるという例は少なくありません。そこはどこかのナンバーで一元化しておくと、また管理もすごく楽かなと。楽というよりも、どこか寄せる宛所がないと困るのですが、やはり違う資料がたくさん出てくる中で、これがどれが代表名なのかを決めるときが結構苦労の種となっていて、胃に悪い作業ですので、そういう胃痛を軽減してくれるようなところがあると有難いですね。

（関野） ありがとうございます。そうこうしているうちに、残りがあと 5 分となってしましたので、そろそろこのディスカッションを締める方向でいきたいと思います。最後に、今回、機構の外からお越しいただいたお三方に、今日のディスカッションや他の方のご発表などもお聞きになられて、やはり人間文化研究機構にはこれをやってもらわないと困るとか、こうしてほしいという要望を、一言ずつ、お伺いしようと思います。では、榎原先生からお願いしましょうか。

（榎原） 今日は皆さんのお話を聞いていると共通するところをおっしゃっているように思います。菊池先生がおっしゃっていたハブというのも、恐らく、データを集めてということではなくて、ここを通過することによって、お互いをつないでいこうということなのだろうと思います。nihuINT も一つの通過点でいいと思うのですが、いったんここを通過することによって、あちこちにつながっていくというような仕組みを考えていければいいのではないかと思います。

それから、さきほど「意外な発見」ということが話題になりましたが、今日、田中先生のお話を聞いて、『古事類苑』へのアクセス数が最近上がってきているというのは、まさに「意外な発見」なのではないかと思いました。恐らく一般の人たちは『古事類苑』をあまり知らない。研究者でも若い人たちはこの頃はあまり使わなくなってきているようです。ところが、検索中、どこかの入口から入ってきて、あれこれと検索しているう

ちに『古事類苑』の存在に気が付いた人が増えている、それでアクセス数が増えているということなのだろうと思いました。これは「意外な発見」の最たるものではないか、忘れられかけていた本が復活してきたというのはとてもいいことだと思います。

（田中） われわれの要望としては、機構にはたくさん貴重な資料があって、実は、データベースを作成、運営している我々のような会社をはじめ、連携したいところはたくさんあるのです。でも、させてくれないのではないかなというのは実は思ってしまう。先ほど申し上げたように著作権の問題などもあったりするのかなと、先回りして尻込みしてしまったり。このあたり、もしウェルカムだよということで言っていただければ、多分われわれだけではなくて、いろいろなところが声を掛けるのではないかと思います。何か、若干やはり声を掛けづらいというか、閉鎖とまでは言わないですけれども、そういうふうに見えてしまうところがあるので、こういった壁がなくなると、もっと官民、いろいろなことができるのではないかと、今日のお話を聞いていて思いました。

（菊池） 今日の研究会でしゃべっているような内容というのは、とてもホットで、かつ重要テーマばかりだったと思うのです。ですので、こんな小規模ではなくて、もっと、それこそこれを可視化していく感じでやっていただくと、今おっしゃっていただいたような、いろいろな外部との連携もできるでしょうし、いろいろ進み出すのではないかなと思います。ありがとうございました。

（関野） どうもありがとうございました。それでは、最後に、永村先生からご挨拶を頂きたいと思います。

（永村） ご挨拶というよりは感想のような所見を幾つか述べさせていただきます。データの利用技術についてしばしば議論になりますが、私は固陋（ころう）な人間 ですので、いまだに活用よりも、データの生成が価値だと思っていまして、史料を調査しデータベースを構築すること自体が、研究者の重要な任務と確信しているのです。しかし自分の研究に目を向ければ、日常的に大正新脩大蔵経のデータベースを活用し、大変に重宝させていただいていまして、このデータベースが仏教に関わる研究のスタイルを変える大きな起爆剤になったことは確かです。その一方で、大正蔵データベースには収められることのない、諸寺に伝わる膨大な聖教類が研究資料として不可欠であり、それらを活用するためにも史料調査とデータベース化が必要とされます。自分の研究という立場から考えて、データを作り続けることが価値であり、その作業を続けることが自分の任務という意識を持ち続けているのです。

さて、nihuINT 利用の問題ですが、その利用率は決して満足いくものではないのですが、その理由の一つにデータの「質」の問題があると思っています。データの「質」というのは、データの精度の良し悪しということではなく、データ群に研究成果に裏打ちされた網羅性があるか否かということです。例えば、東京大学史料編纂所で構築されるデータベースは、明治時代より蓄積されてきた全国的なデータ群によって、時代的・地域的な網羅性がある程度保証されます。ですから検索結果に一定の信頼性をおくことが

できるわけです。また歴史民俗博物館には水木本の古文書群や田中本の聖教群が架蔵されますが、水木本にはその親となる東大寺文書、田中本には醍醐寺聖教のデータベースがあり、それらが一体になったならば、史料群としての網羅性が保証されます。このようにデータの「質」を踏まえてデータベースを構築することが、研究者の基本的な役割であり、研究者が有効に関わり得るのはそのような場面だと思います。

今日のご報告のなかで、榎原先生のお話にあったように、データ群相互のつながりがデータベースの使いやすさに大きく影響するわけです。また、田中先生がご報告になるということで、先月、急きょジャパンナレッジの会員に再度なりました。ジャパンナレッジが提供する語句・用語のデータベースがもつ網羅性と、各機関の研究者が作るデータベースとの相互補完が、データ利用という側面での重要な課題だと思っています。その意味では、先ほども菊池先生が指摘された幅広い利用の喚起も重要な課題となるわけです。

ご挨拶にならない感想を述べましたが、30年ほど前にも京都大学大型計算機センターが中心になって人文学関係の情報処理に関する研究会が開催され、理科系・文科系を越えて様々な研究分野の方々が熱っぽく報告・議論をされていました。その熱気をもう一度復活できないものでしょうか。そして、その場をつくる側として人間文化研究機構の役割もあると思っております。これからもどうぞよろしくお願いいたします。

（関野） ありがとうございました。それでは、ここでディスカッションを締めたいと思います。司会にマイクをお返ししたいと思います。

（大内） これで全てのプログラムは終了でございます。今日はどうもありがとうございました。機構に対して、またnihuINTに対して、いろいろなご提案やご注文を頂いて、これからもう一回、それらを参考にしながら、より皆さんに使っていただけるようなnihuINT、また利用していただけるような機構になっていけるようにと思っております。今日はどうもありがとうございました。これで終わらせていただきます（拍手）。

人間文化研究情報資源共有化研究会報告集　9

Proceedings of the Study on Information Resources of the Human Science Vol.9

令和元年６月20日発行

編集・発行　大学共同利用機関法人　人間文化研究機構
総合情報発信センター
高度連携情報技術委員会
〒105-0001　東京都港区虎ノ門４－３－13
ヒューリック神谷町ビル2F
Tel　03－6402－9200
URL：https://www.nihu.jp/

リサイクル適性Ⓐ
この印刷物は、印刷用の紙へ
リサイクルできます。