KU-ORCASの目指してきたもの
――併せて今後の展望

内　田　慶　市

What KU-ORCAS has been aiming for
— and its future prospects

UCHIDA Keiichi

我在这次小论里首先要介绍现在的东亚文献的数字化的现状，然后说明我们的关西大学开放式亚洲文化研究中心（KU-ORCAS）的到现在的情况和将来的展望等。最后还要提到数字化的各种问题和研究图书馆的构想等。

キーワード：KU-ORCAS、東アジア文献資料、デジタルアーカイブ、
デジタルヒューマニティーズ、秘蔵は死蔵なり、研究図書館

はじめに

私達のKU-ORCASも2022年3月末をもって一応の区切りを迎えることになった。2017年に文科省私立大学研究ブランディング事業に採択され、3年目で思いがけなく梯子を下ろされたが、幸い、その後も法人、大学執行部の支援により最後までプロジェクトを続けることができた。

さて、今や、理系の学問は言うまでもなく、社会科学、人文科学を問わず、デジタル・アーカイブは「当たり前」のものとなっている。KU-ORCASはこうした時代の流れの中で発足し、目指したものはまさにそのような「オープン・プラットフォームが切り拓く新しい人文知の未来」ということであった。

今回は東アジア関係文献のデジタルアーカイブの状況を中心に、それに関わる幾つかの問題について今後のORCASのあり方を視野にまとめておきたいと思う。

1. 世界の東アジア関係デジタルアーカイブの現状

現在世界では例えば広くオープンなサイトとして以下のようなものがよく知られている。

Internet Acrchive

HatchTrust Digital Library

GoogleBooks

Openlibrary

これらは誰でも使えるものであるが、少し専門性に傾いたものとしても、筆者がよく利用するものには次のものがある。

Crossasia（ドイツ国家図書館）

Gallica（BN）

Münchener DigitalisierungsZentrum Digitale Bibliothek（MDZ=ミュンヘンデジタルセンター＝バイエルン州立図書館、IIIF 対応）

DIGITALISERTE SAMMULUNGEN（ベルリン州立図書館）

DigitalVaticanLibrary（DVL=DIGVATLIB）

イエズス会文書館（Archivum Romanum Societatis Iesu）

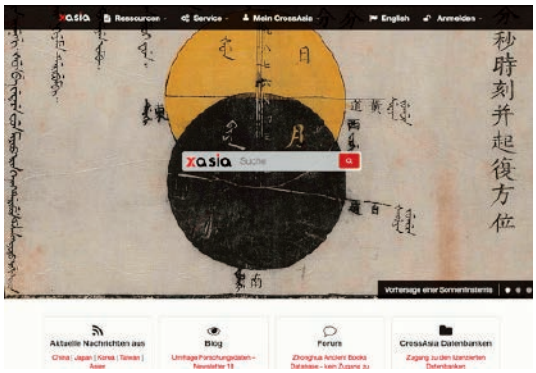
National Library of Australia

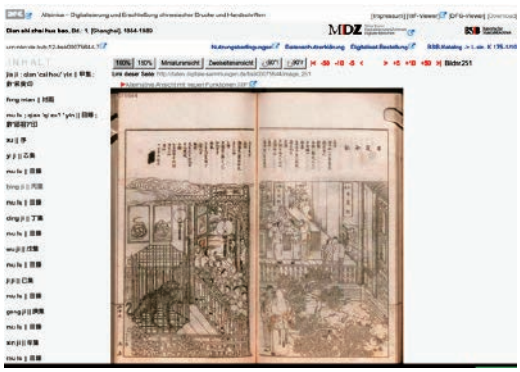
また、大学図書館でも東アジア関係で充実したものには、以下のものがある。

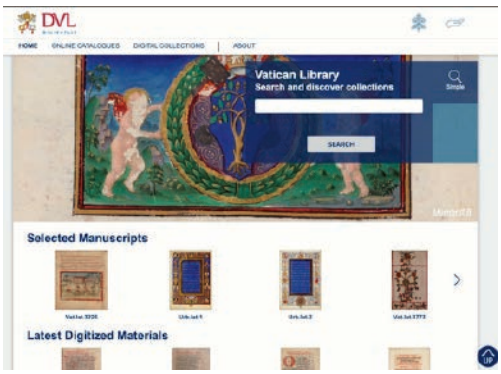
Harvard University（Yanjing）

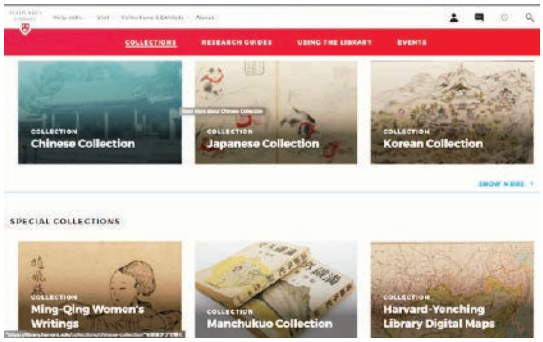
Serica（Bodleian Lib）

Heidelberg 大学（COE）









東アジアのデジタルアーカイブに特化して言えば、この面で間違いなく世界をリードしてきたのは台湾中央研究院であるが、例えば、歴史言語研究所の漢籍電子文献資料庫はこの 30 年間で歴代典籍 1349 種、文字数は 7 億 5 千万余字にも上るデータを蓄積しており、しかも、全文検索可能という優れものである。

中国では中国工程院と浙江大学が中心となって CADAL（China Academic Didital Associative Library ＝中国数字図書館国際合作計画）という国家プロジェクトが進行中であり、既に 700 万冊のデジタル化が行われて公開されている。日本でもかつて、関西大学や愛知大学、それに国会図書館関西館が当初参画を考えていたが著作権等の関係で実現せず、最近、東京大学大学院人文社会系研究科次世代人文学開発センター（下田正弘先生）がようやく日本で初めて加盟している。

この他、環太平洋研究図書館連合（PRRLA）というのもあるが、日本では現在まで関西大学と東北大学が参画しているに過ぎない。

日本国内については、主なものとして以下のようなものがある。

国文学研究資料館の各プロジェクト

国立国会図書館デジタルコレクション

早稲田大学図書館古典籍総合データベース

京都大学貴重資料デジタルアーカイブス

なお、最近は単に書籍のデジタルアーカイブのみならず、以下のように教育研究にも広く利用可能ないわゆる「研究リソース」をまとめたポータルサイトも充実している。

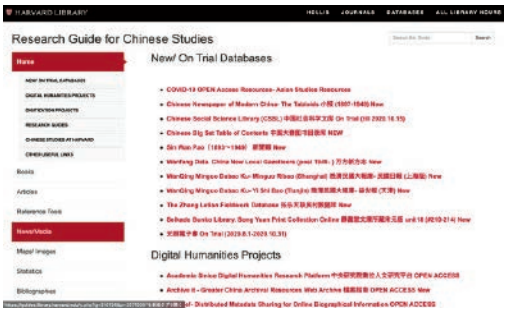
Research Guide for Chinese Studies（ハーバード大学）

東アジア DH ツール & リソース情報（https://www.zotero.org/groups/2337485/dhku-orcas/items/collectionKey/SN63NQVA）

Digital Resources and Projects on East Asia（http://prcurtis.com/DH/resources/）

「国会図書館リサーチナビ：アジアの文化機関のデジタルアーカイブ（中国・台湾・韓国）」（https://rnavi.ndl.go.jp/research_guide/entry/theme-asia-129.php）

北京外国語大学歴史学院のサイト（http://www.globalhistory.cn/islib/index.htm）



中国語に特化したサイトについては、上述の台湾中央研究院の漢籍電子文献資料庫の他に、代表的なものとして以下のものがある。

CTEXT（中国哲学書電子化計画，Donald Sturgeon）

CCL（北京大学中国語言学中心）

BCC（大数据与语言教育研究所）

KU-ORCAS 近代漢語語料庫（関西大学デジタルアーカイブの一部，これに関しては後述）

2. 関西大学デジタルアーカイブ

我々の（KU-ORCAS では世界のアジア研究における「ハブ」的研究拠点として、特に東アジア文献資料のデジタルアーカイブをその主要な目的としているが、現在公開しているコンテンツとしては以下のものがある。

関西大学東アジアデジタルアーカイブ
大坂（大阪）画壇デジタルアーカイブ
泊園文庫デジタルアーカイブ
泊園印章デジタルアーカイブ
アジアの映画関連資料アーカイブ
古文書・古記録デジタルアーカイブ
関西大学デジタルアーカイブ ANNEX（コロナアーカイブ）
近代漢語文献データベース
新規公開コンテンツ
近代漢語文献データベース（近代漢語文献數據庫）

近代漢語文献データベース（近代漢語文獻數據庫）

これらのコンテンツの中心的位置を占めるのは「関西大学東アジアデジタルアーカイブ」であるが、これは関西大学総合図書館所蔵の個人文庫をそのターゲットとしたものである。

関西大学総合図書館には世界の東アジア学研究者の垂涎の的である個人文庫として、次のようなものがある。

内藤文庫（33500 点）…… 内藤湖南（漢籍）
長澤文庫（30497 点）…… 長澤規矩也（国漢籍）
中村文庫（33491 点）…… 中村幸彦（国文）
増田文庫（16184 点）…… 増田渉（魯迅、西学東漸）
吉田文庫（2479 点）　…… 吉田伊三郎（アジア外交）
鬼洞文庫（10309 点）…… 出口神暁（国文）
泊園文庫（16954 点）…… 藤澤東畡・南岳・黄鵠・黄坡

現在これらの文庫の約 7000 点程度のデジタル化が終了し、そのうち公開されているのは約 5000 点となっている。

なお、東アジア文献資料には、書籍以外のものも含まれており、以下のような非典籍資料のデジタル化も行われて公開されている。

漢封泥のデジタル化（20 件程度）
内藤湖南の書簡が 8000 件程度：撮影終了（未公開）
泊園関係の印章類が 300 点程度
内藤湖南文庫所蔵の軸物や貴重書庫収蔵の軸物 700 点程度

また、「近代漢語文献データベース」は近代における西洋人の中国語資料を中心にしたコーパスで、全文検索、複数テキスト横断検索、異体字検索が可能であり、その結果はテキストと画像が連動して表示されるようになっている。

3. KU-ORCAS 文献デジタル化の基本ポリシー＝私蔵（秘蔵）は死蔵なり

これは私がこれまで様々な場所で機会ある毎に述べてきた書籍に関する「持論」であるが、例えば、10 年ほど前に次のようなことを述べていた。

> 研究者の道を志して以来、世界各地の図書館を渡り歩いてきたが、いつも感じてきたことがある。それは日本あるいは広く東アジアの図書館と欧米の図書館との決定的な違いである。そして最近では、これは実は「東西の文化」や「ものの見方・考え方」の違いではないかと思うようになってきている。それは何かと言えば、とことん「利用者」の側に立ったサービスを心がける図書館と、あくまでも「（図書の）管理者」の側に立ったそれとの違いである。

たとえば、これはこれまでにも色んなところで言ってきたことである（「秘蔵は死蔵なり―図書館と文献公開のあり方」『東方』360号、2011. 2）が、貴重書（Rare Book、善本書）の取り扱い方である。日本では（我が関西大学でも同様であるが）、大概が予め（1週間ほど前に）閲覧予約・申請をし、館長の決裁がおりて初めて閲覧が可能となる。また、その複写（duplicate）なども結構面倒で、時には許諾されない場合もある。ところが、欧米の図書館では（私は何十年も利用してきているが）そうした経験をしたためしがないのだ。もちろん、閲覧申請が午前だけとか、午後の何時までとか、幾つかの小さな制限はあるのだが、当日行って申請しても殆ど見せてもらえるのである。

貴重書専門のハーバードのホートンライブラリーでも、大英でも、パリ国立図書館でも、オックスフォードのボドリアンやケンブリッジのワイリーコレクション、ウェードコレクションでも、ローマやナポリ、スペインの国立図書館でもどこでもそうなのである。いつだったか、突然訪れたボドリアンやウェードコレクションでは書庫まで入れてくれたりした。しかしながら、日本や中国では、まずこうはいかないのである。

この違いは一体何なのか。おそらくそれは、「利用者の便」を最優先するということと、「図書の使命とは何か」に対する考え方の相違である。つまり、欧米では「図書はあくまでも人に読まれるべきもの」という考え方が根本にあるのだと思う。面白いのは、それが「著作権」や「所蔵権」などのいわゆる「知的所有権」に対する意識と反比例しているということである。欧米では、この「知的所有権」には極めて敏感であり、その権利を守ることが最低のモラルとなっている。大いに公開するが、他方、その権利は保証するという一見矛盾する考えが徹底しているのである。これが東アジアになると逆転するのだ。「コピー天国」は必ずしも中国の専売特許ではなく日本だってその傾向は昔からあるのだ。特に知的所有権についてはその意識は明らかに低いと思われる。大切なものは人にはなるだけ見せず、他人のものは自分のものとして知らぬ顔。東西の図書館の根本的な違いはここに由来すると私は考えている。

しかしながら、情報公開や国際ネットワーク、文献アーカイブス、データベース化はすさまじい勢いで発展しており、もはや「秘蔵」などと言っているような時代ではなくなっているのである。隠しておいて何になるのか。紙もいずれは朽ち果てて、結局はその使命を果たさずに終わるのである。「図書館は博物館ではない」というハーバード大学イェンチン図書館のJames Cheng館長の言葉を特に東アジアの図書館関係者は心して聞くべきであるし、電子化が一方では、人類の「知の保存」でもあることを認識すべきなのである。

貴重書の扱い方だけでなく、東アジアの図書館が西に学ぶべきことは他にも沢山ある。

WiFiなどの通信インフラやラーニング・コモンズなどは欧米ではもう当たり前、1年365日とは言わないまでも休館日はほとんどなく、しかも夜遅くまで（中には24時間）利用が可能である。バッグを持ったまま書庫に入れ、しかも、そこにはそれぞれに机と椅子が用意されていて、自分が読みたい本もキープしておけるのだ。日本でもそのような図

書館がいつの日か出現することを私は夢見ているのである。（「巻頭感　私の夢見る図書館像―開かれた図書館をめざして」『関西大学図書館フォーラム』第 17 号、2012）

この考えは現在まで全く変わることはないが、これがすなわち KU-ORCAS のポリシーともなっている。

従って、KU-ORCAS では、「研究リソースのオープン化」を掲げ、本学のデジタル化資料のインターネット公開を事業計画として掲げているが、KU-ORCAS が公開する「関西大学デジタルアーカイブ」においては、上記のポリシーのもと、メタデータとコンテンツデータに対して CC ライセンスを付与することで、オープンデータとして提供を行うものとしている。具体的には、メタデータに対しては CC0 を、コンテンツデータに対しては PDM を付し、ユーザーは（著作者人格権を侵害しない限りにおいて）私的利用、商用利用を問わずあらゆる目的において利用することを可能にしているのである。（ただし、これはあくまでも KU-ORCAS で公開するものに限られており、図書館独自のデジタルアーカイブとは別の基準となる）

4. 新しい研究方法の芽生え－デジタル・ヒューマニティズ

さて、デジタルアーカイブの最終目的は単なるデジタル化、アーカイブ化ではなく、それはあくまでも手段であって、新しい研究方法と結びつけて考えるべきであること、すなわち、デジタルヒューマニティ＝新しい人文知の構築を目的とすべきである。

例えば、語彙研究などでは、カードの枚数が勝負という時代があった。つまり、作品を読みながら一枚一枚カードを取っていくわけだが、読んだ作品の量が研究を左右したのである。しかしながら、ある語彙の「初出」という場合、あくまでもその人が読んだものに依拠するわけで、結局はその範囲内でしかものが言えないというわけであった。しかし、今はそうではない。コーパスやコンコーダンスが多く作られ、全文検索も可能となっており、こうした環境の変化は語彙研究を別の次元に押し上げることになっている。N-Gram Viewer やテキストマイニングなどを活用し、更にはそのことで、他の研究領域（概念史、思想史等）との「協同」も可能になってきているのである。

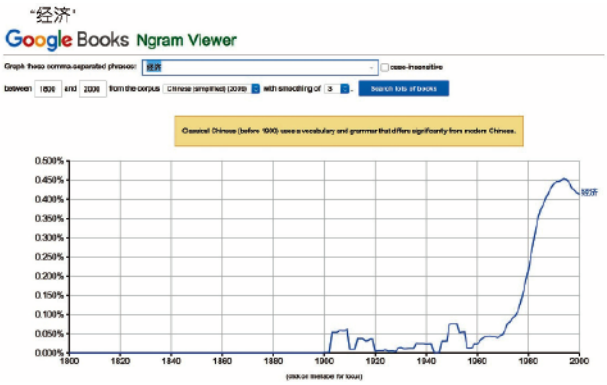
TEI（Text Encoding Initiative）なども最近注目されており、版本間の校勘なども容易に行われるようになった。

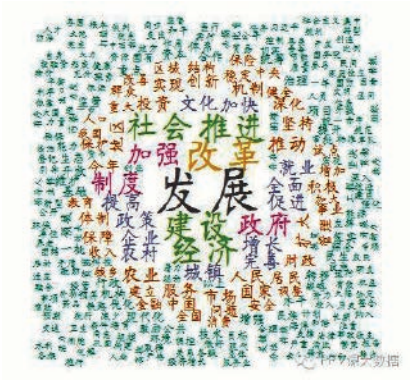
ただ、中国語や日本語などのかつて「2 バイト言語」と呼ばれた言語の処理はやはり欧米の言語ほど簡単ではない。

コンコーダンスやコーパスにしても、その後ろにある電子テキストの入力の場合、今でも大きな障害になっているのが「単語」の切れ目である。これが実は厄介なのだ。また、中国語の場合は漢字の種類の問題もある。簡体字、繁体字、異体字等をどう処理するかである。多くの電子ブックが世に出ているが、単に、現代語や作品講読ならばそれほど気にする必要もないが、

語史研究と言う場合になれば話は別である。字体の違いは時代を反映するからである。これを意味が同じからと同じ漢字で表したら研究にはならなくなる。

こうした問題はあるものの、デジタルアーカイブは確実に新しい研究様式の変化をもたらしたと言えるだろう。





5. デジタル化における諸問題

先ず「サイロ問題」について述べておく。

現在、多くの機関でそれぞれにデジタル・アーカイブス化が進められている。それ自体は好ましい事であるが、しかしながら、一方では横の連携が希薄であり、各所で同じようなものが別々に作られて「閉じた」状態という状況が見られる。

自分たちのサイトに貴重な資料が沢山保管されていても、それが内向きで広く利用されなけ

れば、宝の持ち腐れであるし、また、他の機関でも同じような資料をデジタル化しているのは極めて不経済なことでもある。しかも、アクセスが少なければ、早晩閉じなければいけない状況に追い込まれることになる。

こうした問題を永崎研宣氏等は「サイロ問題」と呼んでいるが、この問題は一刻も早く解決されなければならない。

すなわち、「一国（組織）主義」の打破であり、国内外ネットワークの構築が求められるのである。

また、こうした国際的ネットワークの際に、例えば、画像フォーマット等の国際的な統一規格も求められている。そこで期待されている一つが、IIIF である。

IIIF（International Image Interoperability Framework）とはスタンフォード大学等によって提唱された国際的な Web コンテンツ共有（画像相互運用）の枠組みであり、その機能としては以下のようなことが挙げられる。

同じビューワー内に他機関の IIIF 対応資料画像を同時に見ることが可能。

アノテーション（いわゆるメモ、注釈）機能を追加することも可能。

このようなことから、資料の比較研究が容易となるほか、多様な立場の人の様々な解釈を共有できるようになるメリットが生まれてくるのである。

すでに、スタンフォード大学図書館、大英図書館、フランス国立図書館、オックスフォード大ボドリアン図書館、バチカン図書館、ハーバー大学等等 600 以上の期間が採用しており，日本でも東京大学大学院人文社会系研究科次世代人文学開発センター人文情報学拠点、京都大学図書館機構、関西大学アジア・オープン・リサーチセンターがコンソーシアムに参加している（IIIF を利用するだけならばコンソーシアムへの参加は不要である）。

次に、文献資料のデジタル化はそれが「最終目的」であるか否かの問題である。確かに、デジタルアーカイブの推進は今や時代の潮流であるが、しかしながら、それだけでおしまいではない。それは、実は、上述の新しい研究方法や文献利用のあり方、図書館のあり方と深く結び合ったものでなければならないと私は考えている。

例えば、私たちのKU-ORCASでは「サイロからコンテンツを解き放す」ということから、3つのオープン化を謳っている。

① デジタルアーカイブの構築・公開による研究リソースのオープン化

② アーカイブ構築に関わる研究組織を内外に開く研究グループのオープン化

③ デジタルアーカイブの構築とその活用手法に関わるノウハウや課題を共有し協議する研究ノウハウのオープン化

また、単なる研究者だけでなく以下のような様々なステークフォルダーを想定した。

第一群……専門研究者、専門学会、MLA

第二群……異分野研究者、異分野学会、教育機関（小中高）、企業、自治体

第三群……在学生、保護者、卒業生、受験生、海外からの留学希望者、市民

つまり、本学図書館が所蔵するあらゆる資料群や東アジア文化研究の成果を、知識基盤社会に適合する新たな人文知として、世界に広く提供することを最終目的としているが、これからのMLAに求められるのはこのようなことではないかと考えるのである。

6. もう一つのサイロ問題と将来のあり方

「サイロ化」という問題は実は国内外の連携のみならず、一つの大学や研究組織の中でも存在する。関西大学でも、デジタルアーカイブはKU-ORCASという研究センターが中心になって推進してきており、そのデータ管理もセンター内で行っているが、一方で図書館でもこれまでにデジタル化を行った部分もあり、それは図書館で管理しているし、リポジトリーなども図書館の管轄である。しかしながら、大学全体のITや情報管理はいわゆる「ITセンター」の職掌でありながら、文献資料のデジタルアーカイブに関しては関知しない。従って、サーバーもそれぞれ独自のものとなっている。その更新もそれぞれで行っており、経費の工面にそれぞれが頭を悩ましている。私が関係してきたものでも、学術フロンティア事業、私立大学戦略的研究基盤形成事業、GCOEや私立大学研究ブランディング事業とこの15年余り継続して時限的な外部資金に頼るしかなく、それが途切れたときが「縁の切れ目」となるという状況である。こうした状況は極めて不経済であり、これらを統一して管理する機構というのが必要になると考えている。また、それを支えるものとして図書館での研究部門の設置やサブジェクト・ライブラリアンといった図書館員の専門職化も日本では今後積極的に進められるべきであると考えている。勿論、その場合、MLAの強固な連携確立も求められるはずである。

ところで、実は上掲の過去の文章の中で以下のようなことも述べていた。

ついでに私が「夢想」していることをもう 1 点述べておきたい。それは「アジア学研究図書館」の開設である。関西大学は GCOE や卓越した大学院採択に象徴されるように、東アジア学の研究においては世界をリードする位置にある。それは、東西学術研究所の伝統を受け継ぐものでもあるが、更に、この分野における蔵書において関西大学図書館は世界に誇るべきものを有している。内藤文庫、泊園文庫、長澤文庫、増田文庫、吉田文庫、中村文庫等々の個人文庫はまさに世界中の研究者の垂涎の的となっている。これらの財産を活かすべく、アジア学に特化した「研究図書館」を総合図書館に付設できないかと考えている。そして、「アジア学研究なら関西大学の図書館で」といった、いわゆるアジア学研究の「ハブ図書館」として世界にアピールするのである。そのためには、国際ネットワークの充実が不可欠である。現在、ハーバード大学イェンチン図書館、ルーヴェン大学図書館とはすでに相互利用に関する協定を締結しているが、今後、こうした協定図書館を増やしていきたいと考えている。（同上）

現在日本ではこうした「研究図書館」は東大の U-PARL のみだが、今後、各大学図書館でも検討されていい課題であると考えている。

ともあれ、KU-ORCAS の第 1 期は 2022 年 3 月をもって終了する。その後はいかなる展開を見せるか注目しているが、関西大学の東アジア研究が今後も連綿と続けられていくことだけは疑いないはずである。（了）