

(2014年11月7日開催の2014年度国際図書館協力シンポジウムの音声を文章化した。日本語訳。#数字は英文と対応)

## Zhejiang University Libraries and CADAL project: Retrospection & Anticipation

Huang Chen 黄晨 (浙江大学副図書館長 CADAL 事務局長)

### 『浙江大学とCADALプロジェクト ー省察と展望ー』

ありがとうございます。皆さん、こんにちは。昨晚、プレゼンテーションを少し変更したので、既に印刷されている配付資料にずれがあります。変更を加えたのは、私の話をより深くご理解いただきたいと思ったからです。

#2

CADALのバックグラウンド、その省察、今後の展望を3部構成で説明して参ります。

#3～#8

まず、浙江大学を紹介します。1897年に求是書院(Qiushi Academy)という名前で設立されました。中国の総合大学でトップの三つに入る大学です。総面積は533ha、キャンパスは五つあります。フルタイム学生数4万、教員スタッフ8000名。

図書館は1897年に設立。分館は六つあり、カレッジライブラリーは20を超えています。

(これはいろいろなキャンパスにある図書館の写真です。)

浙江大学図書館の蔵書は、ほぼ全ての学問分野を網羅しており、600万冊を超えています。年間の回転率は100万冊です。

図書館のウェブサイトですらに詳しい情報をご覧いただくことができます。

#9

CADAL(China Academic Digital Associative Library)プロジェクトについて。これは中国デジタル図書館国際協力計画(中国学術デジタル共同図書館計画)と呼ばれています。浙江大学がスタートさせたプロジェクトなので、オーガナイザーとしての役割をずっと続けており、浙江大学図書館がこのプロジェクトの推進センターになっています。これには、浙江大学副学長がこのプロジェクトの発起人の1人であるという経緯があります。CADALプロジェクトはすぐれた技術と豊富なデジタルリソースを使って学術的なデジタル・ライブラリーを構築することを目的としています。多くの学術分野・カテゴリーにおいて、多様な言語により国内外の図書館・学術機関と研究者たちにサービス

を提供することを目的としています。

#### #10～#11

プロジェクトは 2001 年に開始され、三段階で進んできました。

節目となる重要な出来事をいくつか紹介いたします。このプロジェクトは 2001 年に開始されました。米国では「万卷の書プロジェクト」、中国では CADAL として知られています。2006 年の時点でデジタル化した図書は 100 万冊、これが第一段階です。2012 年の時点では、250 万冊の図書がデジタル化されています。

#### #12

この起こりは、2001 年に中国とアメリカのコンピュータ科学者が協力して、「中米百万冊デジタル図書館プロジェクト」に着手したことです。これは 100 万冊のコレクションをインターネットを介して誰でも無料で読め、検索もできるようにしようという試みです。(これは)発起人としてプロジェクトを開始されたおふた方の写真です。1 人が中国工程院会員で浙江大学・前副学長の Pan Yunhe (潘云鶴) 教授、もう 1 人が米国のカーネギーメロン大学の教授である Raj Reddy 博士です。

#### #13

2002 年 9 月、中国国家開発改革委員会教育局は、5 ヶ年 10 期計画で大学における教育・研究の情報支援をするため CADAL を実施することを提言しました。皆さんは北京大学で行われている CALIS プロジェクトをご存じでしょう。CADAL と CALIS のプロジェクトが合わさって、China Academic Digital Library and Information System の枠組みを構築しました。

#### #14～#18

第 1 段階では、中国のトップ 16 のほとんどの大学がプロジェクトに参加しました。

第 2 段階では、国内からの参加が増えて 70 となりました。

海外からも続々と参加しています。

(これが)プロジェクトのこれまでの成果のいくつかです。世界最大で専門のデジタル化センターが深圳経済自由特区に設立され、40 を超える大学にデジタル化センターが設置されています。デジタル化能力が 1 カ月当たり 2100 万ページ、7 万冊となります。これはたいへんなデジタル化能力です。

また、中国各地に 8 つのデータセンターと 33 のサービスセンターを建設し、24 時間サービスを提供しています。

#### #19～#20

CADAL は既に 270 万点のデジタル化を終え、中国における最大規模のデジタル・ライブラリーとなりました。

オンラインリソースは、180 万点です。

#21

(これは)2013 年の CADAL ポータルへのアクセス数の統計です。1 日当たり閲覧・50 万ページ、ダウンロード・26 万ページ超。アクセス ID で数えた正味 1 日当たりの蔵書検索数 184,000 件、そのうち 47%が公衆ネットワークの利用者です。ポータルのアドレスは <http://www.cadal.zju.edu.cn> です。

#22

2013 年には 1,900 以上の国立大学が CADAL を利用しています。

#23

海外からは 500 の学校が CADAL ポータルを訪れてています。1位がアメリカで、2位が日本です。

#24

(これが)トップ 20 大学における CADAL ポータルの利用数です。944 万点を超えています。中には、CADAL ポータルの利用が自分の大学の利用の半分に相当する大学もあります。

#25

CADAL と外国の機関との協力関係を紹介します。2012 年 11 月 9 日、ドイツ・ベルリン州立図書館と協定を結び、データベースとネットワーク・プラット・フォームを介して中国の歴史と文化を共同研究することになりました。

#26

2012 年 11 月 25 日、日本国立国会図書館副館長が 5 名の専門家とともに CADAL を訪問しました。

#27

昨年の初め、インターネット・アーカイブと契約を交わし、70 万冊の英語の書物の寄贈を受けました。また、デジタルによる 400 万冊の貸し出し合意が成立し、七つのコンテナに梱包された 15 万冊の英語の図書が既に大学に到着しています。米国の非営利組織であるインターネット・アーカイブは、世界最大のデジタル・ライブラリーを運営している組織です。

#28

2013 年 3 月 27 日、CADAL 地域協力センターが香港に開設されました。香港城市大学で開所式

が行われ、工程院副院長の Pan Yunhe (潘云鶴) 院士と香港城市大学学長の Guo Wei (郭位) 教授が出席されました。

#### #29～#30

2013 年 4 月、オックスフォード大学中国研究図書館館長・中国コレクション長であるデビッド・ヘリウェル氏の訪問を受けました。

2013 年 5 月、米国研究図書館協会副会長とカリフォルニア・デジタル・ライブラリー副館長、そしてハワイ大学の Diane Perushek 氏が CADAL を訪問しました。

#### #31

プロジェクトのハイライトは次のとおりです。

- CADAL Open API
- カスタマイズ可能なおすすめ機能
- 「章」単位でのデジタル貸し出し
- デジタルライブラリーからナレッジセンターへ

#### #32

私たちは CADAL 参加者のために Open API を開発しました。参加館はこの Open API を使って CADAL のリソースを自館のウェブパック(OPAC)に統合することができます。ユーザーは自館の蔵書検索ができるだけでなく、CADAL ブックシェルフや CADAL 電子ディスクの検索もできます。

#### #33～#35

(これが)カスタマイズされた CADAL ポータルです。

カスタマイズ機能により、ユーザーは検索・借り出し・注釈付けをすることができる仕組みになっています。

(これが)おすすめ図書です。登録ユーザーは CADAL の中に個別のポータルを持っており、CADAL にアクセスするとおすすめ図書が出てきます。

#### #36

その他のハイライト。

CADAL のデジタル・リソースの 60%はまだ著作権保護の対象です。そのため、ユーザーや協力機関からよく聞かれることは、どうすれば著作権で保護されているものを閲覧に供せるのかということです。解決策としてチャプター・ベース(章単位)のデジタル貸し出しモデルを構築しました。

#### #37

(これらが)CADAL デジタル貸し出しモデルの構造とフローです。サッと CADAL ブック・シェルフを

開いて何ページかすると、このパターンが見つかります。

#38

ユーザーは「章」単位で貸り出すことになります。

#39

ある人に「章」が貸し出されていると、他の人が同時に借りることはできなくなります。現在は 70 の国内大学図書館が参加しているので、書籍によっては全体で 40 冊ぐらい所蔵されています。ですから、ユーザーが「章」借りをしている時は、同時に 40 冊分、他のユーザーも借り出すことができます。本の現物は図書館に置いたままです。

#40

借り出した「章」は14日間利用可能です。14 日経過すると自動的に返却されますし、ユーザーが期日前に自分で返却することも可能です。

#41

CADAL プロジェクトの成果により、デジタル・ライブラリーの技術が中国でどんどん広がり、普及してきています。2009 年には、教育局デジタル・ライブラリー・リサーチ・センターの設立が承認されました。

#42

2011 年末に、CADAL の 10 周年記念式典が浙江大学で開かれました。発起人である Raj Reddy 教授と Pan Yunhe (潘云鶴) 院士が基調講演を行いました。その講演において、デジタル・ライブラリーはスマート・ライブラリーに形を変え、最終的にはナレッジ・センターに進化するだろうと述べています。

#43～#44

2012 年、ユネスコ事務局長のイリナ・ボコヴァ女史が CADAL を訪問しました。

同年、ユネスコ事務局長補のグレッチェン・カロンジ女史も CADAL を訪問しました。そして昨年夏、ユネスコは CADAL の成功を受けて、中国エンジニアリング・ナレッジ・センターを創設することを決議しました。

#45

以上 CADAL プロジェクトについて簡単にご紹介しました。ここからはわれわれのプロジェクトの将来計画についてお話をします。

#46

実際のところ学術情報の世界ではオープン・アクセス、オープン・ナレッジ、そしてオープン・イノベーションが一般的な考え方です。この3つの考えをもとに CADAL プロジェクトの第三段階を展開しています。

#47

クラウド・テクノロジー、インテグレイションを使いたいと思っています。現在はビッグ・データが着目されているのでビッグ・サービスを契約すべきです。ユーザーと双方向に知識を共有し、情報関連の研究所などと透明性のある形でユーザーを共有し、図書館間で実験成果を共有してゆきたいと思います。伝統的なデジタル・ライブラリーの構造の中では、資源と実験結果を図書館間で共有してゆきたいと思います。

ユーザーのためにデジタル化を進めたいと思いますが、今では、ユーザー側でもますます多くのオンライン・マクロ・コンテンツを持つようになってきました。そこでユーザー側のマクロ・コンテンツを CADAL で収集・組織して、他のユーザーにも使えるようにして知識の共有を図りたいと思います。

「ユーザーは一つのウェブサイト、あるいは自分の所属する図書館のウェブサイトを訪問するだけで、世界中のコンテンツを見られるようにする」「ユーザーは CADAL や他のプロジェクトのことを知らなくとも、たったひとつのポータルを訪れるだけで世界中の情報資源を見つけられるようにする」、そのためには CADAL は次の4つのパーツを作らなければなりません。情報資源、技術、サービスそして相互協力関係です。

#48

構造化される以前の膨大なデータを処理することは主要な課題のひとつです。大量データは海のような性質をもつので、「データの大洋」と呼ばれていますが、ビッグデータという呼び方のほうが有名です。私たちは「データの大洋」という考え方を4年前に提唱しました。その後「ビッグ・データ」として知られるようになりました。

#49

ナレッジ・ベースやマイクロ・コンテンツ、リソースの構造をオープンにしたいと考えています。これらのオープンな構造により、伝統的なデジタル・ライブラリーの構造を B2B(組織→組織)から C2B(消費者→組織)のモデルに変革させる、言いかえると、ユーザーが自分の図書館や情報関連機関をデザインできるということです。

#50

(これが)知識ベース共同構築の流れです。どの大学、機関、個人でも、CADAL への参加申請は次のようになります。

応募者はCADALのデジタル・リソースと自分のデジタルリソースを比較分析し、共同構築計画を

申請します。CADAL では応募された素材を評価して、リソース構築委員会で検討します。申請が受理されると申請者は CADAL メタデータ・プラットフォームを使って書誌メタデータを作らなければなりません。次に CADAL から電子書籍を入手し、デジタル化を完了させます。

応募者は CADAL にフルテキストを提出するか、機関リポジトリ上で CADAL メンバーがフルテキスト・サービスを受けられるようにするかを選択できます。別の言い方をしますと、応募者はフルテキストを保持したままで、メタデータだけを CADAL ポータルに提供することもできます。

CADAL の編成指針と統合情報検索機能により、高い透明性をもって共同知識ベースが CADAL に構築されていきます。

#### #51～#52

(これがその例です。)

これは浙江大学が作った碑文データベースです。浙江大学より CADAL に提案がありました。「浙江大学は碑文データベースを CADAL に提供しするので、CADAL には碑文に関するすべての素材、本、刊行物、論文をこのデータベースに付け加えてほしい」とのことでした。そうすることでオープン・ナレッジベース上に全ての情報が結集されて、ユーザーに提供することができるようになります。

#### #53

マイクロ・コンテンツ共同構築ですが、登録ユーザーの全員がポータルで自分のマイクロコンテンツをつくることができます。ラベル、インデックス、ブックレビュー、お薦め、利用者ノート、注釈付けなどが可能です。利用者はコンテンツを公開することも非公開とすることもできます。公開を選択すると、管理者が内容を点検し、その後その書籍の一部として公開されます。非公開を選ぶと、本人だけが見ることができます。

#### #54

(ご覧のように)このデータベースでは、ユーザーはある碑文の画像について、「もっと鮮明な写真や古い写真を持っている」とコメントをつけることができます。また、自分の持っている画像を提供することもできるし、このデータベースのメタデータに追加・修正を加えることができます。

#### #55

次はリソース共同構築についてです。漢字 OCR は精密性に欠け、これを利用した文献デジタル化は困難、というのは皆さんご存じのとおりです。ですから、ユーザーに OCR で変換されたものを読んでもらって修正を加える必要があります。皆さんにテキストをお渡ししますので、校正をお願いできれば助かります。

CADAL は、ユーザーの貢献度に応じて、全文検索・原本調査などのインセンティブを与えています。

#56

これは別の例です。過去数十年にわたり中華民国の図書と逐次刊行物をデジタル化してきましたが、その時代の新聞は何ひとつデジタル化できていませんでした。そんな時、ある会社を見つけました。その会社は中華民国の新聞のデジタル化に特化しているので、我々は自分たちの素材をひとまとめにした。そうするとユーザーはこのポータルから中華民国に関するすべての素材を調べることができるようになります。書籍や定期刊行物ならば、CADAL のユーザーは CADAL ポータルから全文をダウンロードして随時読めますが、上述の新聞が読みたいという場合はその会社のお世話になることとなります。

#57

ご覧のように新聞や定期刊行物、書籍などの全ての材料を一つのポータルから検索することができます。新聞のフルテキストを読むこともできるし、新聞の画像を入手することもできます。

#58

次は個人ひとりひとりのためのユビキタスサービスについて。CADALのクラウドサービスから進化を遂げた個人用クラウドという形で、自分ひとりのためのインターネット図書館を作ることができます。

CADALは登録ユーザーのために電腦空間を提供するので、ユーザーはそれを使ってオンライン書齋を作ることができます。ユーザーはインターネットを介して自分のインターネット図書館にアクセスできます。PCでもよいし、モバイル・タブレットを使ってもよいです。PCにもモバイルタブレットにも対応したインターフェースが開発済みです。簡単にCADALの知的資源にアクセスしてマルチディスプレイに映し出すことができます。

#59

CADAL の目標はユーザー中心のデジタル・ライブラリーを作ることです。我々はユーザーたちの知的資源の獲得方法や管理方法、創作活動に注目しています。目的は、ユーザーが知的探求ができるように足場を提供することです。具体的には次のようなことです。

私たちはユーザーたちの活動である検索・貸し出し・蔵書・購入・お薦め・本の交換に注目しています。伝統的な図書館は利用者の検索と図書の貸し出しだけを見ています。

(私たちは決して彼らの購買活動を気にしているわけではありませんが、)このデジタル・ライブラリーのインターフェイスにおいては、ユーザーたちは本を中心に、あるいは人を中心にして自分の全知的活動を管理できます。本屋で本を買う、オンラインで注文する、他のデジタル・ライブラリーから、あるいはCADALブックシェルフからダウンロードする、図書館から本を借りるなどいろいろありますが、自分で蔵書管理ができます。ブックレビューを管理したり、図書リストを作ったり、ラベルを作ったり、ブログを書いたりできます。



以上のように一括りにすることで、ユーザーはまるで自分の書斎にいるようにデジタル・ライブラリー・ポータルを研究・調査に利用することができます。

#### #60～#61

携帯用アプリケーションを作りましたので、例として紹介します。簡単なインターフェースです。図のカメラマークをクリックしたり、本のカバーのバーコードをスキャンすると、自分が持っている本や買って来た本、借りて来た本など、情報を管理する画面に自動的に移ります。全ての情報をみんないっしょにすることができます。さらにこのインターフェースでは、本のバーコードをスキャンすると、書籍情報を取得することができます。これはユーザーの所属する図書館が、探している図書を所蔵しているかどうかを調べるシステムです。もし自分の図書館が所蔵していないのならば、下にあるボタンをクリックすると購入希望が出せます。CADAL が持っている本とわかれば、このボタンを押すと借りることができます。ユーザー自身の図書館にその本の現物があることがわかれば携帯電話で予約できます。

本の価格を調べることもできます。Amazon や中国 dangdang.com などのオンラインで購入することができます。

#### #62

このシステムでは、友だちに自分の本を貸すこともできます。カメラ・マークのボタンをクリックして、自分のデータベースに貸した人の名前を入れられます。そうすれば誰が本を借りたのか思い出させてくれます。また、本の分類も可能です。ラベルを付けて、蔵書を自動的に分類することができます。

#### #63

国際協力として、もっと多くの海外の大学・機関に、クラウド・ソーシング・モデルに参画していただき、デジタル資源・技術・サービス及び人的分野で共同開発を行い、成果を共有したいと思いません。

#### #64

(これは参加申請の) 三つのモデルのうちの一つですが、ドイツ・ベルリン国家図書館のように全デジタル・リソースのコピーを CADAL に提供する方法があります。彼らは、CADAL に電子化した中国の図書を提供し、CADAL も彼らにコレクションを提供しています。

また、CADAL に書籍を提供して、デジタル化してもらう方法もあります。図書館の中には、特別コレクションを CADAL にデジタル化してもらったうえで他の図書館と共有しているところもあります。

香港のようにコンソーシアムを作って CADAL のプロジェクトに参加し、かなりの数のデジタル・リソースと書籍を CADAL に提供している例があります。

以上のような三つのモデルにおいて、CADAL は IP レンジの範囲内でデジタル・リソース・サービ

スに対するアクセス権限を利用者に付与しています。

#65

以上で発表を終わります。ありがとうございました。