

馬場毅（東亜同文書院大学記念センター）：鈴木先生の質問は確か北アフリカのマグレブ以南に同じようなことをやったかどうかということでしたよね。鈴木先生、直接英語で話されたほうがいいと私は思いますが。

Suzuki: I would like to... My question mentioned about several places in the Middle East. There is an institution for studying those areas. So I would like to compare with those kinds of institutions. Is it possible to compare with the TDS activities or not? That is my question.

Ms. Bastid-Bruguière: Well, the French institutions in the Middle East, all the schools which were the French schools, for instance there is a secondary school for long time in Istanbul, also in Cairo, these institutions were directed to training the local elite, not training the French. Occasionally French residents would send their children there but they would pursue studies to enter French university, not to stay there. So there was that kind of institution. Then there was another kind of institution, for instance, L' institut Français du Cairo. There is the same in Damask, then the École Française d'Extrême-Orient and even the... These institutions are research institutions, it is true, but usually most of them, or generally it was geared toward the past - old Japan or literature - not trade. I mean, generally it was not trade. It was antiquity. It was humanities. It was not practical and that is a big difference with the TDS which is practical, I think. Later on, for instance, recently the [inaudible] has a program on Japanese economic development or Japanese financial things, so now it is geared to

study of present day. But it is very academic. Not practical.

馬場毅（東亜同文書院大学記念センター）：鈴木先生、これから休憩に入りますので、この後直接お話しください。ブルギエール先生、どうもありがとうございます。それではこれから15分間休憩に入ります。したがって3時22分から再開したいと思います。ちょっとご案内をします。一つはこの館、今下りてきましたが、東亜同文書院のDVDをここで上映させていただきます。またお茶が外の廊下のほうにありますので、ご自由にお飲みください。懇親会ですが、藤田先生が最初にお話しされていましたが、無料ですのでふって多数ご参加いただきたいと思います。以上です。休憩します。

(休憩)

馬場毅（東亜同文書院大学記念センター）：それでは時間になりましたので再開します。休憩後の最初のご報告はミシガン大学のニキ・ケンジ先生にお願いしています。最初に簡単にニキ先生の履歴を紹介させていただきます。ニキ先生は上智大学文学部をご卒業後、1977年に渡米されました。セントジョーンズ大学でアジア学、中国近代史専攻の修士号を取得されました。その後プラットインスティテュートで情報学の修士号を取得されています。現在ミシガン大学アジア図書館ライブラリアンとしてご活躍されています。今日のご報告の題は「ミシガン大学における東亜同文書院およびアジア系文献史資料のデジタル化とその利用」ということです。それではよろしく申し上げます。

ニキ・ケンジ（ミシガン大学）：私の名前はニキ・ケンジと申します。よろしく申し上げます。本日は藤田先生からお招きを受けてかなり自分自身の

中で受けていいのかちょっと苦しみました。というのは、私はお二方のような学者ではなくて、資料を扱う人間なので、資料という面からしか東亜同文書院を見ることができなかつたのです。したがって、海外における東亜同文書院云々という場合のカテゴリーにあてはまるのだろうかということをもとに感じました。それで藤田先生にお聞きしましたら、自分が思ったようなことをやっていただいて、そのほかにミシガン大学の今グーグルの先端をいっていることを述べていただけませんか、ということでしたので、お引き受けすることにしました。

実は愛知大学とはずいぶん昔から関係がありまして、ミシガン大学のアジア図書館に私が就任して10年になるのですが、その半ばごろに東亜同文書院の大旅行史マイクロフィルム、また調査旅行報告書マイクロフィルム、この二つの大きなセットを国際交流基金のオープンコンペティションで2年かけてプロポーザルを書いて、その二つのもを私どもほうにいただくことができたのです。したがって日本以外のところにこの資料があるのは北米ではもちろんミシガン大学だけです。ですから、そういう意味から私は愛知大学との交流を深めていったわけです。

2005年だったと思いますが、こちらのライブラリーに招かれて、そのときにやはりこのような立場でものを申したことがあります。そのときには北米におけるアジア学の動向について述べたことを記憶しています。今回はそれに今の先端の、これからのライブラリーの姿、あるいは学者、あるいは学者になろうとしている皆さんのこれからどのような資料を渉猟しなければいけないか、その形態がどのようになるかについてちょっと述べていきたいと思います。

私どもミシガン大学のOPAC、すなわちライ

ブラリーの中にあるその大学の図書館の中のオンラインリサーチをやります。そのときにキーワードで東亜同文書院と入れますと、ミシガン大学のOPACは数を6000と出してくれます。それだけたくさんさんの東亜同文書院に関する日本語だけではないすべての言語の資料が存在していることを表すわけです。でもこれはキーワードサーチですから、それがすべて別個の本かどうかということはまた別の問題です。ただ、東亜同文書院というワードを入れますと約6000のヒットが出てくると。したがって、何冊本を持っているかということに関して私はまだ詳しく調べたことはありません。

とにかく今日は、東亜同文書院は私にとっては私の修士課程の勉強が中国現代史に集中したものですから、そのときに私は東亜同文書院の存在を知ったわけです。それも私は日本でではなくてアメリカでそれを知ったのです。そしてライブラリアンになりまして、コロンビア大学に就職しました。それが1983年です。

1983年というのは全世界の中でかつてのカードをめくったライブラリーのシステムがようやくコンピュータを使ったシステムになりかわろうとしていたときです。それはチャイニーズ、ジャパニーズ、コリアンというCJKラングイッジが初めて1台のコンピュータでライブラリーのインターフェイスに入力できると。そういうシステムをアメリカが作り上げたわけです。

そのハンズオンのときからの仕事を私は1983年からコロンビア大学で始めました。したがって今は世界中にCJKという言葉が氾濫し、CJKという言葉でサーチができて、すべてのインターフェイスはCJKラングイッジが目前面に出てくると。そういうことはあのときには考えられもしなかつたことなのです。私たちが入力するときにはこんな大きなキーボードでいろいろな偏とか画と

かいろんなものを入れながら漢字を作成していく。それでコンピュータの中におち込んでいくという操作をやったわけです。私はそのときからずっとコロンビアの書庫にあるいろいろな書物を渉獵しているときに東亜同文書院のものがやはりありました。

ですから、私の人生の中で、要するにインテリクチュアル・インスティテューションのアカデミック・ライブラリアンはどのような姿勢をもたなければいけないか自問自答したときに、もちろんヒューマニティとソーシャルサイエンス、この二つの轍で広く浅くというのが我々図書館員の一つの使命なのです。けれども、先ほどのお二方の先生たちは非常に深く、学者としての見事な対比を私たちは学んだわけです。でも私の立場から東亜同文書院を見るときには、やはりそれは資料としてどのような価値があり、どのようにして学者に対してそれをサービスできるか。あるいは大学の学生さんに対してそれをサービスできるか。そのところに意識がいくわけです。ですから、ずいぶんと違った見方をしてしまうのです。

私もここに招いていただくために東亜同文書院に関するいろいろなことをやはり少しは深く知りたいたい。それで、いろいろなものを渉獵して読み始めたのですが、ここにいらっしゃる先生方のものを読み始めたときに、俺は大変なことになったと思いました。レイノルズ先生の論文や藤田先生とか栗田先生の本などは大変深く拝読しました。しかし、私はブルギール先生のはフランス語が主なので…すみません、読めませんでした。とにかくそういう意味で、私自身勉強しなくては、と思っているのが現状です。

私はまず、ライブラリアンの姿勢として東亜同文書院は自分のライブラリアンとしての人生での一つの轍と決めました。それがコロンビア大学で

す。また、そのときにもう一つの轍を日本の博学の細菌学の天才である南方熊楠、この二つの轍を私は自分の通常8時半から5時まで、たいていは8時から7時ごろまでの毎日の仕事の中で二つのものを自分の意識の中でどうやってやればいいのかと思って30年やってきたわけです。

それからミシガン大学に移ってから10年ですが、ミシガン大学が今なぜ騒がれているのかというところで東亜同文書院と今お別れをします。そして、ミシガン大学が今世界に一生懸命宣伝しようとしているグーグルのプロジェクトについて述べていきたいと思います。

こちらにパワーポイントのスクリーンが出ますから、それに沿って日本語でそれが出てきますので、皆さんはそれを読んでいただければいいと思います。グーグルとミシガン大学の馴れ初めは、元々グーグルの副社長がミシガン大学のMBAを取った人だと聞いています。それでミシガン大学に声がかかって、世界でいちばん最初にすべての大学の図書数百万冊をデジ化するという目的でグーグルとミシガン大学は契約を結んだわけです。その後ミシガン大学のすべての図書館のグーグル化が始まりました。でも、ミシガン大学はそこにあるように、ミシガン大学そのものがグーグルと手を結ぶ前にミシガン大学の図書館だけでデジ化を進めていたのです。

そのデジ化を進めていたプロジェクトの名前はMBooksと言います。ミシガン大学のMとBookを一つの言葉にしてMBooksと言われます。今ここに出ているのは、すべてミシガン大学にこれだけライブラリーがあります。あの小さな田舎町のキャンパスの中に今ナンバーで記されているのが全部ライブラリーです。私が働いているのはその本部のHatcher Graduate Libraryです。

今そこにありますすべてのライブラリーがほとんどのものをすでにデジ化が終わっています。ただ、サイズがあわなかったり、あるいは状態が悪かったりとか、いろんな意味でそれを少しずつ改良しながら入れているので、まだ全部終わったとは言えないのですが。たとえば私が働いていますアジアライブラリーは合計81万冊を超えています。その81万冊のチャイニーズ、ジャパニーズ、コリアンラングイッジのものはほとんど終了しています。これに関わってくるものはもちろんコピーライトです。これがいちばん大きな問題で、それを外へ出せるかどうかということは問題になります。しかしデジ化は内部ですでに終了しています。

今日本ではこれと同じような形になっているのは慶應大学です。慶應大学一校だけこの方法でデジ化を進めています。大きな問題はやはり抱えていると思われれます。グーグルは今、このときのOne Million Digitized Booksとありますが、それは100万冊が終わった時点での情報をここにしました。これは2008年2月で100万冊が終了しています。こういう専門的なことは私にもあまりわかりません。だからどのようなキロバイトだとか、そういうことはあまり問題にはならないので、現在200万冊をこの時点では終了しています。私が来る前に200万冊を終えていますから、今は220~230万冊終わっていると思われれます。

MBooks についてですが、2008年ミシガン大学が自身のアクセスシステム MBooks を立ち上げました。その後初めてグーグルとの契約があって、MBooks がもっていたデジ化された情報をグーグルは受け取って、それからデジ化されていない図書館に入ってきて、それをデジ化していくという形でミシガン大学のグーグル化が始まったわけです。

ミシガン大学はグーグルとあまりにも密接に

co していたのですが、その間にいろいろな軋轢がありまして、これではやはり良くないと。そしてグーグルが100%ミシガンに自分たちがデジ化したものは無料でミシガンのものとして、グーグルが持つもの、ミシガン大学が持つもの、というふうにしてこのように二つに分けたデジ化の結果をシェアします。ですから、はっきりとミシガン大学は自分のところだけ持ったデジ化の記録があります。グーグルはグーグルで同じものを持っています。

ここでミシガンは MBooks から Hathi Trust という一つの大きなグループのコンソーシムと言いますか…コンソーシムを形づくって、カリフォルニアのシステムの大学と、それからミシガンの属しているシカゴも入るミッドウエストの辺りの大学が共同で Hathi Trust という一つのグループを形づくって、そこで自分たちがデジ化されたものの共同使用をやるのではないかということになって、ミシガンがその中心になりました。今 Hathi Trust がかなり大きく取りざたされて、Hathi Trust が今持っているすべての記録の70%はミシガン大学の記録です。あとの40%がカリフォルニアステートの十いくつの大学のもの、あと五大湖を中心としたあの辺りの大学のもの、この二つのグループが一緒になって Hathi Trust を作っています。いちばん最後のほうにいろいろな URL のサイトが出ますから、そこでお調べになればこれに対する詳しいことがわかると思います。

次に、ちょっと文字が読みにくいですが、一応日本語になっています。私自身目が悪いのであまり読めないのですが、これは一つの例として挙げておきます。

その次に、これもちょっと…アクセシビリティというのは非常にわかりにくいのですが、アクセ



スの許容可能範囲と訳されていますが、通常のインターフェースを改良し、よりアクセス可能になると。それはどういうことかという、大学はバリアフリーと言うのですか、いろいろな意味で身体の不自由な方々にもオープンしていますので、目の不自由な方々のためには大きなインターフェースを使った資料が見られるということも Hathi Trust のほうでは改良に改良を重ねて私たち、自分たちだけのそういう方法を使っているということです。

プロジェクトワークフロー、これはもう皆さんがご覧になっても全然面白くないと思います。私もこういうものは面白くないです。そして、デジタル化のための選定と準備、これはバーコードがほとんどのすべてを解決しますので、はじめにどんな古い本でもすべてバーコードをとりつけていきます。それがプロセスされているかどうかということはまったく度外視してバーコードをつけて、そのバーコードでその本の昔の記録があやふやなものでもバーコードをつけていき、そこではじめてグーグルの人々がそれをきちんとした記録として塗り替えていくわけです。

ここでちょっと、グーグルが何百万とある図書館の本を物理的にどのような形にもっていくのか、ちょっとこれを見ていただきたいと思います。グーグルのプロセスがここに図式されていますが、視覚的に面白くないので、まずこのような形で大学の書架を全部空っぽにします。これはどういうことかという、大きなトラックが大学の図書館の裏側につきます。そこで若い職員、これはグーグルの職員で、ダダダーっといくつものカートを持って入ってきます。そして私たちは何も知らないのです。どこで何をしているかわからない。ただ、書庫の中に入ってきて、彼らはその中に全部本を一つ一つ全部積んでいくのです。そのときにはすでにその本はすべてバーコード化され

ています。だから、このように空っぽにする前に若い職員が自分の専門的なラップトップを持ってきて、一つのフロアに幾人かの若い職員が本を全部一つ一つ抜き取っていきます。そしてそれを全部調べて、バーコードがついていないものは全部その場でバーコードをつけていくのです。バーコードをつけたらそれを全部自分のコンピュータに入れていきます。そういう形ですべてのバーコードが終わった段階でこのような書庫を空っぽにしておきます。それは私たちの手の届かないところで行われていて、それがわずか数日の間にまた元のように書庫の中に返ってくるのです。

魔術のようなものですが、巨大な組織がここで動くのです。それはアナーバーの小さな街のどこかのビルにその本を運んで、そこで数百人の若い技術者がそれをスキャンするわけです。スキャンしてそれをすべてグーグルのデジ化の記録の中に収めていき、それをまた元の書庫にこうして返すのです。だから毎日毎日何トンという大きなトラックが裏の入口について、朝から夕方まで本を出し入れして数日前にやられた本が返ってきます。それをまた何十人かが書庫に運んで、自分たちがすべてを元のとおり直すのです。私たちはいっさい何もしません。それも契約のうちの一つだということです。

次に、今後側…右側のほうにトラックが見えますね。このようにしてカートに全部入れてトラックで運び去り、この裏口からまた運び込むのです。毎朝毎夜こういうことが繰り返されます。そして見事なチームワークでグーグルはやるのですが、どのようなところでどのような形でやっているかは企業秘密で、決して我々には知らされません。ミシガン大学の図書館の中で知っているのはただ一人です。その一人の副館長だけがグーグルと渡り合って、グーグルとの会議ですべてを決めます。ですから私たちライブラリアンは何もしないでい

いと。要するにすべてのものはグーグルがやるという形でミシガン大学のものを全部グーグル化しているわけです。

グーグルが不合格品として出してくるものがありまして、これはスキャンできないと。それはどんなものかという状態が酸性してぼろぼろになっているとか、あるいはページが飛んでいるとか、彼らは1枚1枚スキャンをしますから、そういったものはそのときに選んでそれを我々のところに送り返してきます。送り返されてきたものの中で、我々のほうでそれをなんとかしてもう一度デジ化できるかどうかということ調べて、できるものは全部私たちが手でそれをやっています。

さて、グーグルの不合格品は右にあるような形で全部グーグルが突き返してくるわけですが、特にバーコードや書誌の情報がない、あるいは書籍の状態が悪い。そして問題が解決された場合にはもう一度デジ化に踏み切ります。それは先ほど説明したとおりですが。そして、ミシガン大学のみのデジタル化というのがそこで初めて行われます。それはグーグルが突き返してきたものに関してミシガン大学だけがそれを直して、デジ化して、それはミシガン大学の記録であって、グーグルの記録にはしないと。だから Hathi Trust のほうの記録の中にミシガンのものが練り込まれていきます。その数はそれほどたいしたものではないと思います。

さて、日本のグーグルで非常に大きな問題になっていますのが著作権についてです。これは向こうでも著作権について大きな問題になりました。それで著作権の調査としては、1923年から64年の著者不明などのいわゆるグレーゾーンの作品を対象にしてもっと詳しく我々が調べ上げなくては、このことに関しては著作権で訴えられる可能性がある。そういうものがグレーゾーンのものは特に

多いのです。1923年以前のものに関してはコピーライトがありません。したがって、1923年以前のものは何の許可も必要なくどんどんデジ化していくということです。著作権データベースにまずいろいろな情報を保存します。それから IP アドレスでフィルタをかけていくと。私も専門的なことはどのようにするかちょっとつかまえてはいないのですが、出版社がもうそのときには存在しないと。そういうことが昔の本には多々あるわけです。そういうときにはどのようにすればいいか。グーグルはそういう面倒くさいことはやらずに突き返してくるわけです。ですから、それはミシガン大学で懇切丁寧の一つ一つをクリアしていくということになります。

グーグルと著作権との間の係争の問題が次に出てきます。これは品質調査でどのような形で、要するにぼろぼろになったものとか、あるいはデフォルメしてしまった出版物をテイクケアするか。これは大変努力のいる仕事ですが、根気よく続けています。Hathi Trust 対グーグルブックの検索、要するに今ははっきりと我々が持っています Hathi Trust というグループのミシガン大学を中心にしたグループの共有の財産のデジ化されたものと、グーグルをグーグルのブック検索はどう違うか。左側にありますのが Hathi の考え方で、全文 PDF による全体ダウンロードはなし、そして検索のみ。ごめんなさい、それで検索結果はキーワードが含まれている箇所すべてに対して。これと同じ対照的に右側にいきますと…これはグーグルとなっていますがグーグルの間違いです。グーグルブック検索の考え方は PDF による全体ダウンロード付き、またスニペット、各検索キーワードに3スニペット、これは非常に小さな情報をそのまま我々はとって、それが次の段階に進むことの一つのギアになるということです。それで左側が我々のやり方で、右側がグーグルブックのやり方なのです。それぞれに一つの欠陥、それぞれに

良い点があります。

著作権のクリアに関して。追加コンテンツを加える能力、あるいは長期保存へのコミットメント、イリノイ大学はもうすでに学術的ニーズに対応して著作権のクリアを自分たちがやり始めたのです。著作権のクリアくらい面倒くさいことはないと思います。

将来はどのようになるかということですが、Hathi Trust はより大きくより良くと。他の電子コレクションをこれからどんどん複合していく、飲み込んでいく。そしてサービスがよく行き届く。そのグループに入っていればお金を払わないですべてのものが読めるという形のものにサービスをしていく。それに対抗するグーグルに関しては、グーグルがデジ化したものに関しては、グーグルはあくまでもプロフィットオーガニゼーションですから、ユーザーはそれに対してお金を払っていくという形になると思われま

さて、そのいちばん下にありますエスプレッソ・ブック・マシンやその他の POD オプションと書いてありますが、このエスプレッソ・ブック・マシンに関しては、次の別のパワーポイントの説明でしたほうがいいと思います。今はその部分を頭の中に入れておいてください。グーグルとどのような係争があったかということです。これは2005年、米国作家協会による訴訟が起き、2008年10月にグーグルは和解を発表しました。2009年9月までその和解は延長されています。詳しくは黄色のところにありますように、このサイトに皆さんが行っていただけると、いろいろな情報がすべてそこに書かれています。そして、個人的な質問も、たとえば皆さんがインターネットで、Eメールで質問しても答えてくれます。それはもう絶対に答えてくれます。ただし、それは英語で質問していただきたいと思いますが。

次にスキャンするもの、これは今色で分けてあります。黄色はすべてのものがスキャンされています。緑色が2009年から2010年にスキャンするもの、白いものに関しては、これは要するに??のライブラリーに関しては、すべてのものがビデオとDVDなのです。ですから、そのものはスキャンできないということになります。今のところ、マイクロフィルムとかマイクロフィッシュのコレクションで私どもの持っているものも、それをデイブリケートすることはありません。ですから、東亜同文書院の二つの大きなマイクロフィルムはスキャンされていません。というのが今の状態です。

さて、このオンデマンド印刷出版の可能性というのがあるのですが、これはグーグルの大きな流れの中で、ミシガン大学が別個の段階でオンデマンドの出版の可能性として作り上げてきたものがあります。そのことについてはこれからの問題になります。特に出版関係の方に関して、あるいは大学の教職員の方の要望に関して、これはある意味で革命的な一つの方法なのです。ですから、たとえば皆さんが今ライブラリーに行って本を借ります、本を読みます、返します。あるいは本屋さんに行って本を買います。自分の書棚に入れます。こういう今までの出版と本当の関わりがこの段階で非常に大きく崩れる可能性が出てきたのです。

オンデマンドということは要望の多い出版物で、神田なんかのマーケットにはもうないという本をなんとかして再販したいというのがオンデマンドの根本的なものですが、オンデマンドに関してはコピーライトとかやはりいろいろそういうことが関わってきます。ただ、そのオンデマンドは、それが出版されるとかなり値段としては高いのです。ここにも出版社の方が来ていますからわかると思うのですが。ところがそのオンデマンド版でしかも異常に安価に手に入る方法が作られたわけ

です。

これはヨーロッパを中心にして作られたものですが、今全世界にエスプレッソ・ブック・マシンというのが27台あります。これは非常に高価な機械なのでそれほどたくさんは今のところは作られていません。このエスプレッソ・ブック・マシンというのはコーヒーのエスプレッソを作るのと同じような意味で、一つの目的に向かってぎゅっとコーヒーのエッセンスを搾り取るような形なのでエスプレッソ・マシンと名づけたのですが。

今アメリカの例をとりますと、1923年以前の本に関してはもうコピーライトはないわけです。したがってHathi Trustにしても、グーグルは別にして我々の持っているデジ化されたライブラリー、Hathi Trustの1923年以前のもの、学者はそういうものをどうしても必要とします。博士課程、あるいは修士課程の学生さんが自分の分野でどうしても1923年以前のものがほしいのだけれどもマーケットにはそんなものはもうないと。あったとしたら、それは古本屋に行って高いお金を払って買わなくてはいけない。そのようなものをどうやったら安価に、しかも遠いところまで買いにいかなくてもすむか。そういう要望が多いわけです。特にそのフィールドにおける学者が必要とする場合、早急に必要なのです。

その場合に、Hathi Trustの200万ある記録の中に1923年以前に出版されたものです。すでに全文こちらにある場合に、先生はそこでしめたと思うわけです。先生はライブラリーのところに行って登録して、私はこの本を読みたい、しかしどこにも売っていない、そしてこの本を自分のものとしたいという要望をします。そうするとライブラリアンはそれに対してコンピュータで調べるとこれはもうコピーライトも何もない。しかも私たちはこの記録を持っている。それでは、ということでその

エスプレッソ・マシンにかけます。すると、本当にすごい勢いでエスプレッソ・マシンが動くのです。動いたらドッドッとプリントしていくのです。こんな本があつという間に目の前に出てくるのです。そして、それは自動的にパジネーションも皆ついて、上の本の体裁のカバーも全部つけるのです。それはこんな色をした、非常に普通の紙の、もちろんハードカバーではない、そういう紙がぐるりと全体をしてちゃんと綴じてあるわけです。

それがもしもこんな太い本であれば、それは10ドルです。このエスプレッソ・マシンで作り上げた本マキシマムのお金が10ドルです。そして8ドル、6ドル、4ドル、3ドルとその量によってお金が変わります。でも、それを先生は1923年以前のもを目の前でわずか10分ほど、先生はどこかで講義して帰ってきました、もうできています、と。それでお金を払ったら、そこでは枚数を勘定していますから、これは6ドルですと言えば、先生は6ドル払えばその貴重な、マーケットにない本を手に入れることができるのです。これが、これからの皆さんの、特に大学の出版の、要するに大学の学者、あるいはリサーチャー、そういう人たちが本当に手が出るほどほしいという形の出版物として脚光を浴びてきているわけです。

今全世界に27台しかないエスプレッソ・マシンの1台がミシガン大学にあります。もう1台はトロント大学にあります。北米にある大学の中ではこの2校がエスプレッソ・マシンを動かしています。それ以外に、私にはどこのパブリックライブラリーかは記憶にないのですが、カリフォルニアのどこかのパブリックライブラリーが1台持っている。それ以外はほとんどヨーロッパです。そして多くは、皆さんここに来ていらっしゃるパブリッシャーと同じように、要するに出版社がそのマシンを買って、その要望に応えようとしているわけです。

ですから、これからの皆さんの、要するに私たちの知的な欲望を満たす材料として、インターネットを使ったグーグルのインフォメーション、PDFで出てくるインフォメーション、そういったものは私のように目の悪い人間には非常に読みにくいわけです。1枚1枚オーダーしてはお金を払ってかえていくわけです。私にとっては非常に面倒くさいと思うのですが、それはないよりマシで。すべてのものがデジ化されたときには皆さんはそうやって読めるわけです。しかし、やはり学者は…こういうことを言って非常に悪いのですが、私もそのうちの一人ですが、非常にエゴの強い人間です。だから、そういったものは自分のものとしたい。自分のところに置いておきたい。そういう意味では4ドルとか6ドルでもうマーケットにもないものを手に入れることができる。これはやはりこれからの知的産業を牛耳る人間としては福音になるのではないのでしょうか。ですから、こういう意味で私はこのオンデマンドの印刷のことをグーグルと話して皆さんにお伝えしたいのですが、グーグルは、グーグルということとデジ化されたインフォメーションがそのライブラリーにあれば、そのライブラリーのデジ化の記録を呼び出して本にすることができる。もちろんそれはコピーライトと関わりますが、そのことをちょっと皆さんにお知らせしておきたかったのです。

さて、ミシガン大学は結局ステートのお金で経営されています。ですから、私は公務員です。ですから、ライブラリーそのものはお金のやりとりができません。ですから、アマゾンブックストアが私たちの第三セクターとしてお金のやりとりをすると契約を結んで、今そういう形でブックオンデマンドの形でミシガン大学のCenter for Japanese studyの過去パブリッシュした本56冊すべてアマゾンのブックストアを通せば買うことができます。でもそれは普通のブックオンデマンドです。ミシガン大学はただ単なるグーグルをや

っているというだけではなくて、ライブラリーはグーグルが来る前にすでに将来のデジ化、要するに電子化されたインフォメーションを皆に供しようという形でMBooksを始め、それでグーグルが参加して、グーグルとグーグルを別にしてHathi Trustという、我々のグループのインフォメーションのシェアリングの方法をとって今日まできています。

さて、次に、これで一応グーグルに関してはこれからのいろいろなクエスションというのはこういう形で出てきます。ですから、もしも皆さんがこれ以上のことというときには、このHathi Trust ウェブサイトを探されると、この中を通じますといろいろな情報が手に入り、しかもそれに対して懇切丁寧にクエスション&アンサーが入っています。ですから、それを参考にしていただくと大変ありがたいと思います。ましてやミシガン大学の、この最後のほうに書かれていますが、ジュリア・ロベットというところにEメールでお聞きになれば、彼女は非常に懇切丁寧に皆さんに状態を知らせてくれると思います。彼女が私の名前を使ってください、と言って私に渡したのです。ですから、皆さんは自分の質問をEメールで出されると、彼女のほうから必ず回答がまいります。一応簡単ですが、私の今日の報告はこのようなものです。

馬場毅（東亜同文書院大学記念センター）：どうもニキ先生、ありがとうございました。最先端の図書の電子化と言いますか、グーグルと組んだ電子化のお話をお聞きして、実は愛知大学も2012年4月にはささしまに参りますが、そこではスペースの関係で図書を全部持っていけないという問題がありまして、今それをどうするかという問題が大変重要な問題になっていますが、それに対しても大変示唆的なお話だったのではないかと思います。この中には実は図書委員会の方もいらっしゃる

るし、ライブラリアンの方もいらっしゃるし、さらに本屋さんもいらっしゃいますので、どなたでも結構ですのご質問いただければと思います。いかがでしょうか。簡単なご質問。はい、どうぞ。キシマ先生。

木島：愛知大学のキシマと申します。愛知大学の、という以上に今馬場先生から指名を受けました図書館員をやっています。今ニキ先生が発表してくださったことに関して伺いたいことが二つあります。一つはHathiというTrustグループに入ると、著作権が残っているものでも学術研究に特化する、要するに売らないということであれば、Hathi Trustに入っている大学の人はどこかの大学、Trustに入っている大学のものは自由に見られると考えてよろしいのでしょうか。

それからもう1点はこのレジメにもありますが、検索のみという形でPDFのダウンロードはできないと書いていらっしゃいますが、ということは、スキャンして画像として保存するのではなくて、文字化してどこかに置いていらっしゃるのかなと思うのですが。英語の場合はスキャンでヒット率がかなり高いと思うのですが、先ほどの話でいきますとCJKでいくとかなり確率が落ちるのかなという予感ももっています。ただ、英語の場合、もしくはフランス語、ドイツ語もあつたりますと、完全に校正をして一字一句間違えないような文字データを持たせるというのはかなり苦しいことなのかなと思ひまして。OCRをかけてヒット率80%、90%くらいでOKということにしていらっしゃるのか。その辺はある意味企業秘密なのかもしれないのですが、お聞かせいただければと思います。

最後にもう一つは、ではどんどんよそにある本が手元に来るといふことでいきますと、全部万々歳のような気がするのですが、結局ミシガン大学

であれ慶應大学であれ、実物の本を持っているところがやはり生き残る中心になるのかなという印象を、今話を聞いてもちました。ということは、愛知大学は東亜同文書院関係の書物、旅行史であれ省別全史であれ、いろんなものを持っているわけですが、やはり持っているところが中心になってこれから動いていかないといけないのかなという印象もちました。これについては正確なお答えではなくて、やはり持っている人が偉いんですよ、という感触をニキ先生はおもちなのか、それともこれからは持っていなくてもよそから取り寄せられるという世の中が変わっていくのか。その辺の感触も教えていただければと思います。よろしくお願いします。

ニキ・ケンジ（ミシガン大学）：先に最後の質問なのですが、実物を持っているところが強いのではないかとおっしゃったのですが、それはそのとおりだと私も思います。ただ、それだけではなくて、先ほどありましたように、将来、要するに実物はミシガン大学にあります、でもデジ化されたものは共有です、ということですので、それに関してはどこもイコールだと考えられると思います。ただ、ミシガン大学の大学の中に在籍しているものであれば、実物を見ることができます。ですから、もしも何らかの形でデジ化されたものに不具合がある場合は、カリフォルニアの我々のグループの、Hathiのグループがそれをコンプレインしても実物を見ることができないわけなのです。ところがミシガン大学の学生さんや先生たちは、書庫に行けばそれを見ることができます。そういう意味で、実物を持っているところが強いということにはなると思います。

2番目のものですが、ちょっと私は先生のおっしゃっていたことが頭の中であまり像を結ばないのですが、もう一度要点を簡単なことで結構です。



木島：今日いただいたレジメのところに Hathi の特色ということで、グーグルとの対比で、ということだと思のですが、全文（PDF による全体ダウンロードはなしで検索のみ）というふう書いてある…

ニキ・ケンジ（ミシガン大学）：わかりました。

木島：検索をするために文字データはどれくらいの正確さで持っているのかということところです。

ニキ・ケンジ（ミシガン大学）：私はグーグルの専門家ではないので、私がやっているのではないのでそのパーセンテージというのは全然わかりませんが、サーチをする場合にはすべてローマ字でやると思います。ですから、CJK の言葉を使ってサーチする場合には、それがどれだけのヒット率かというのは、私も試したことがないです。ミシガン大学の場合はたとえば日本のものに関してはローマ字を使ってやります。ただ、マーリンというライブラリーの OPAC のシステムの中でサーチする場合は、今は CJK の言葉を使えばヒットします。ですから、それが何十%かということとはわからないのではないのでしょうか。というのは、記録そのものがローマ字だけで書かれているものと、我々の使っているチャイナ、ジャパン、コリアのスク립トが入っているものとあると思うのです。もしもそれで文字が入っていれば、それに対するヒット率は機械がやることですから絶対にヒットします。そう思います。ただキーワードサーチということになれば話は別だと思えます。意味がおわかりでしょうか。

木島：要するにローマ字での世界ではうまくいっているけれど CJK だけの世界では…

ニキ・ケンジ（ミシガン大学）：いや、そういう

ことではなくて、うまくいっているのではなくて、ローマ字でやればすべての…要するにアメリカのすべてのシリリック・ランゲイジというものの情報は全部ローマ字化された情報をまず入力するわけです。それプラス CJK なら CJK の言葉を分類する段階で加味していくわけです。ただ分類する段階で CJK の言葉が入っていない情報もあるのです。したがって、CJK の言葉でサーチした場合には、それが入っていないものはヒットしないということになります。ですから、いちばんいいのはローマ字でサーチすると両方のものが出てくるということです。なぜかという、CJK ランゲイジのものをそのままぶつけてボンとカタログはできないわけです。これは必ずローマ字化されないと。

だから、はっきり言いますとライブラリーコングレスのクラシフィケーションでクラシファイする場合には、まず100のところがローマ字が先です。その次に100のところが漢字、あるいはひらがな、カタカナ、その次また200がローマ字、そして CJK と出てきます。だから常にパラレルで入力しなくちゃいけないのです。今の CJK のカタログは。ただ、それは現代なのですが、その前の情報は CJK がない場合があるのです。ミシガン大学で言えば1986年以降、すべての情報は CJK を入れています。それ以前のもの全部ローマ字です。だから、その以前ものを引っ張りだすときにはローマ字で入力すれば全部出てくるわけです。現代のも出てくるし、過去のも出てくるということです。ただ、グーグルの場合は、グーグルは全部を完璧な状態でやっているはずですから、私はグーグルのほうは外部の人が使う場合、お金を払った場合は、それは全部の情報が出ると思います。ただ、Hathi の場合は私どもの持っている記録ですから、グーグルがグーグルしてくれたものと、グーグルしなかったもの、両方 Hathi は抱えています。だからそのパーセンテージがど

れだけかと言われると私も考えたことはないのですが、よろしいでしょうか。

木島：わかりました。たぶん慶應は考えざるをえないことにはなっていると思いますが、ミシガン大学のシステムそのままではCJK文化圏では簡単にはいかないのかな、という感触を私はもちました。

ニキ・ケンジ（ミシガン大学）：はい。OPACをCJKでサーチすることは非常に簡単で可能性がありますが、その場合にはOPACの中に入っている以前のものはCJKの言葉が付加されていないが、しかしたとえば「夏目漱石の生涯」というのが日本語で入力されていなくてローマ字だけで入力されている時代のものに関しては、漢字とひらがなで「夏目漱石の生涯」と入れても、理屈の上では出てこないと思われま。ただ、1986年以降我々が一生懸命CJKを記録に付加しようとやってきた記録に関しては両方入っていますから、それはたぶん出てくると思います。機械がそれに関しては絶対に間違いを起こさないとはいえませんが。

木島：わかりました。最後の、持っているところが偉いのだということに関しては…

ニキ・ケンジ（ミシガン大学）：木島さんのそのものの言い方が僕にはちょっと納得がいかないのですが。偉いのではなくて、持っているところが偉いというよりも要するに…

木島：いや、私は偉いというのよりも、むしろ持っているところは義務を負っていると考えないといけないのかな、というふうに今日のお話を聞いて少し考えを変えたところ。持っているところが率先して公開の方向に進んでいかないと、持っていないところはしようもないわけですか

ら…

ニキ・ケンジ（ミシガン大学）：そうですね。ですから持っているところは、要するにHathiのすべての、今の時点ですべての何百万とある情報の中で70%のものはミシガン大学のものであるということは、あとの30%がほかのHathiが入れてくれた…だからほかのHathiのグループはまだ全部入れていないです。彼らがもっとどんどん入れてくれれば、このパーセンテージは変わってくると思いますが、今のところミシガン大学のすべてのものを供給、出してしまったので、それでHathi Trustというグループを作り上げたので、今のところ70%の情報はミシガン大学のすべての情報をそこで見ることができる。それ以外に30%、ミシガン大学が持っていないものもそこで見ることができる。だから、ほかのところから言わせれば、30%のものは自分たちのものだけれども、あとの70%の情報はミシガン大学のものだという。でも自分たちが持っている数が少ないですから、ミシガン大学のほうが多いですから、当然それはそうだと思います。ですから、さっきおっしゃったように持っているところが偉いのではなくて、持っているところが供給しなくてはいけないという形になると思います。

木島：同文書院関係のものでいきますと持っているところはやはり愛知大学がいちばん多いのかなと。

ニキ・ケンジ（ミシガン大学）：そうですよ。

木島：そうすると率先して公開をする努めが生じるのかな、というふうに感じたりもしました。

ニキ・ケンジ（ミシガン大学）：絶対にあります。そうしてください、お願いします。



木島：はい。どうもありがとうございました。

馬場毅（東亜同文書院大学記念センター）：それでは時間となりましたので。どうもニキ先生、ありがとうございました。それでは第4報告になりますが、愛知大学の東亜同文書院大学記念センターの武井義和さんのご報告をお願いしたいと思います。最初にちょっと武井さんの略歴をご紹介します。1995年に愛知大学文学部の史学科を卒業されまして、2006年愛知大学大学院博士後期課程を終了されて、博士（中国研究）の学位を習得されています。2006年から現在まで継続していますが、東亜同文書院大学記念センターのポストドクターの職を務められています。主な著作として、まず博士の学位論文として「上海における朝鮮人社会の歴史的考察 1910年から45年」という、ご存知のように朝鮮人は戦前の場合は大日本帝国に所属している日本人でしたので、それが上海にいた場合にどうなるかという大変面白い問題を取り扱っておられます。それから愛知大学の東亜同文書院記念センターオープンリサーチセンターの1号に載っていますが、「東亜同文書院に関する先行研究の回顧と今後の展望」という研究史の整理をされています。同じように研究史の整理で「中国における東亜同文書院研究」、こちらの場合は中国のほうですが、これは愛知大学の国際問題研究所の132号に載っています。以上のように、特に愛知大学の東亜同文書院記念センターに所属して、東亜同文書院大学の研究では若手の注目株ということになるかと思います。それでは今日のご報告の内容は「第2次大戦後の欧米における東亜同文書院研究」ということです。よろしく願います。

武井義和（愛知大学東亜同文書院大学記念センター）：ただいまご紹介にあずかりました武井義和でございます。よろしく願います。私の報告の課題は第2次大戦後の欧米での東亜同文書院に

関する先行研究の整理を行うものですが、その作業を通じて、どのように同文書院を論じてきたのかという点をおおまかながら述べていきたいと思えます。その上で、東亜同文書院を日中両国だけでなく、欧米も関連させてより一層国際的な広がりをもつ研究へと発展させるための分析視覚について自分なりの見解を出してみたいと思えます。言語問題の関係上、アメリカにおける先行研究の整理が中心となりますので、この点ご了承下さい。

さて、私のテーマは「第2次大戦後の欧米における東亜同文書院研究」ですが、最初にお断りしておかなくてはならないことがございます。タイトルに欧米という言葉がついていますので、本当でしたらヨーロッパとアメリカの研究動向を取り上げなくてはなりません。しかし、私が理解できる言語の関係上、また、先行研究を調査して本日まで得られた文献の関係上、アメリカで発表された研究を主に取り上げることとなりますので、その点をご了承ください。なお、英語の論文や図書をいくつかご紹介しますが、その日本語訳についてはこちらのレジメをご覧ください。

私は今までに日本と中国で発表された東亜同文書院に関する研究の整理を行ってきましたが、欧米における研究の整理は初めての試みでした。このシンポジウムを迎える前の時間的余裕が必ずしも十分とは言えない状況の中で調査を進めてまいりましたが、端的に言いまして、日中両国に比べて東亜同文書院を扱った研究は非常に少ないです。もっとも、私は今回東亜同文書院に限定して調査を行いましたので、もう少し裾野を広げて、たとえば中国近現代史、日本近現代史という幅広い観点から見れば、わずかながらでも東亜同文書院が関連的に取り上げられている研究もより多く見い出せたかもしれません。

最初にこの点に関して二つほど事例を挙げた