

## 第 10 講 ジャパンサーチとデジタルアーカイブ活用基盤

高野 明彦（国立情報学研究所名誉教授）

ジャパンサーチは、書籍等分野、文化財分野、メディア芸術分野など、さまざまな分野のデジタルアーカイブと連携して、我が国が保有する多様なコンテンツのメタデータをまとめて検索・閲覧・活用できるプラットフォームである。このジャパンサーチについて理解を深め、基本的な理解を図った後に、メタデータの連携方法について考えます。

### 【学習到達目標】

- ・ ジャパンサーチの目的について説明できる。
- ・ メタデータの連携方法について具体例を挙げて説明できる。

### 1. ジャパンサーチとは

ジャパンサーチというのは国の真ん中辺でやっているサービスなのですが、企画は内閣府の知的財産戦略推進事務局に、推進委員会及び実務者検討委員会というのが置かれまして、大体7年ほど今は活動しています。

ここでデジタルアーカイブという新しい仕掛けを日本の中にどういうふうに根づかせていくのかという議論をしております。私はこの実務者検討委員会というものの座長です。

実際のサービスがジャパンサーチという名前の検索サービスが上がっているのですが、これについてはNDL、国立国会図書館というところが責任を持って運営しているという形になっています。

ただ、国会図書館が美術館、博物館のデータをいろいろ集めたりするというのはなかなかやりにくい面もありますので、直接だとやりにくい面もあるので、この委員会でそういう活動をするということになっています。

国の中のいろんなところでデジタルアーカイブというのが小さいものも大きいものも構築が進んでいるわけですが、それらを一覧できるような場所をつくろうということがまず、第一の目的です。

ポンチ絵を描いてありますけども、アーカイブ機関というのがあって、これは何をアーカイブと捉えるかっていうことによって随分違うのですが、デジタルアーカイブというのにとらわれずに考えていけば、博物館、美術館っていうのは



ジャパンサーチ

大体物をためて置いておく場所ですからアーカイブの機関だろうし、文書館みたいな公文書館に代表されるような文書館はもちろん名前のとおりアーカイブすることが目的の機関なわけです。

あるいは大学とか研究機関っていうのは、一義の目的は教育ですけども、教育の中にはいろんな資料を活用したり、それを集めて初めて分かることについて深めていったりするということが含まれますから、大学の研究を進めていくとそこに何かが集まってくると。これをある人はアーカイブと呼ぶという形なわけです。

これまでアーカイブっていうのはそういう物理的なフィジカルなコレクションっていう場合が多かったわけなので、なかなか遠くからは使いにくくて、研究者っていうのは大体そういう物があるところ、資料があるところを全国あるいは海外まで出向いて、それを調べるっていうのが普通のスタイルだったのですが、これだけデジタル化が進んで、当たり前前にデジタルの技術を使えるということになりましたので、だったら貴重な資料であってもみんなと共有したい、世界と共有しても構わないというようなものであればどんどん公開して、それを相互に利用し合うというような形が理想的というふうになってきております。

それをジャパンサーチでは加速したいというのがもともと願いです。また、片側に活用者がいて、アーカイブ機関というのを、自分のホームのアーカイブもあるかもしれないですが、大体はアウエーといいますか、知り合いを頼ってそういうところにたどり着いて調べさせてくれということでこれまでやってきたと思うんですけども、そういう人間を介さなくても資料と出会う機会というのを提供したいというのがジャパンサーチの目的です。

ですから活用者とそのアーカイブ機関の間に立って、何かうまく探せる仕組みを提供しよう。探すっていうのが、グーグルがはやって以来、何か研究自身は探すことみたいに思われるかもしれないですが、探した先のことをするのが多分活用者にとっての本当に重要なことですから、どこにどんなものがあるよっていうことを一覧できるということが、そういうアイデアを生み出す一つのきっかけにもなるということで、この真ん中をやりたいということです。

しかし、ちょっと考えれば分かるのですが、ここが、アーカイブ機関が例えば1000館ほどがあって、それをじゃあ束ねましょうといいますか、そこにどこがどういうものを持っているのかを総覧できるようにしましょうとなりますと、この真ん中のサービスはその1000館と直に付き合わなきゃいけないことになるわけです。そうすると例えば契約書を一つ交わそうとすると、1000部の組

織とやり取りをするということで、この真ん中のこうサービスを上げるより前に疲弊してしまうようなことが起きます。そこで工夫として考えたのがこのつなぎ役という場所です。すなわち似たような組織は似たような契約をして、似たようなメタデータを集めよう。あるいは集める前に、メタデータの流儀みたいなものがちょっとぶれていたら分野ごとに整理をしていただいて、分野でまとめたものを頂く。そこを一発契約を交わして提供いただくほうが合理的だねということを経パンサーチの立ち上げ時には考えました。

これは我々のアイデアというわけじゃなくて、後で多分、時実さんがお話しされるとおもいますが、ヨーロッパアーナ（Europeana）というヨーロッパを中心に進めているこういう同種の取組があるのですが、そこではアグリゲーターと呼んで、何かを集めてきてアグリゲーションして提供するというような形で、ヨーロッパアーナはアグリゲーターとだけ付き合っていけば、その先の何千という組織とデータをやり取りできるよってというような形をつくられて、割とうまくいっているというのを我々は知っていたので、僕たちもジャパンサーチを上げるときにそれに類したことをやろうということで。

ただ、アグリゲーターという片仮名だということもありますけど、日本語でいうと何なのかわからないですが、何か収奪されるような気がする。アグリゲーションビジネスなどというビジネス書が本屋に行くといっぱい並んでいますけども、並んでいた時期でしたが、大体割とアグレッシブといいますか、えっ、そんな言葉にされちゃうの？というような話が多いので、アーカイブ機関がそれで警戒されては元も子もないので、そうじゃなくて柔らかく日本的に、つなぎ役をするのですという形で名前もつけました。

そこで、この3つの構成、活用者がいて、アーカイブ機関がいて、それをジャパンサーチに取り次いでいくつなぎ役がいる。場合によっては利用者からの要望に応じて、このつなぎ役が整理の仕方なども工夫していくということで、この3つはそれなりに関係づいているという形になります。具体的にはこのスライドは後でダウンロードできるようにすると思いますが、この辺のリンクをたどっていただくと、政府のいろんなものが見えると。ちょっとクリックします。こんな報告書とか、取りまとめサイトみたいなものにたどり着けるとおもいます。こんな名前で引くと、大体7年分の資料が出てくるとおもいます。

次に行きますと、そうやって立ち上げたサイトがあると。僕は割とこういう話をするときはライブデモをする。スライドを一枚も作らずにやるっていうのが好きなのですが、そうやると大体どこかに引っかかって時間を取られて最後までたどり着けないっていうことが多いので、今日はちょっと自制して、最後の



Europeana

最後に少しデモをするかもしれませんが、スライドで頑張っていこうと思っています。

ちょっと調べてみますとジャパンサーチというのは2019年2月に試行公開したのですね。仕組み自身が大体動くのかとか、こんなでこぼこ感のあるデータの集め方でうまく役に立つかしら？ フィードバックを頂こう、ということでこれをやりました。1年半後ぐらいですかね、1年半後ぐらいにちょうどコロナ真っ盛りの夏、8月25日だったかに正式版公開ということにいたしました。ここでデータベースとして、データベースっていても大もあれば小もあるんですけど、何十万も入っているデータベースからほんの数百しか入っていないデータベースまで108ぐらいのところからデータが集まって、メタデータの件数でいうと2000万件を突破するぐらいになったと。2000万件といっても本のカタログだとか、それから公文書館のように、タイトルだけは見せられるけども、公文書の本文は公文書館に来ないと見せられませんというような、そういうものも含めてのものなので、それほどリッチな感じでは当初はなかったですけども、それでも2000万件を突破したっていうのは一つ大きいメルクマールになったということで、ここで正式公開というふうにしました。

公開当日とかは何万も何十万も来たんですけども、大体平均すると6000件ぐらい。もうちょっと行ってもいいかなっていうのが当初感じたところですが。そして2年半ぐらいたっているのですが、昨日調べで見ると197データベースと。データベース、今度はどちらかというと小さいものが増えたんですけども、地域の発信のようなデータベース、今日の多分セミナーの全体のテーマはその辺にあるのかなと思っていますが、地域に埋もれているような資料を整理して上げてくるというような特徴的なものをどんどん入れましょうということで、ここ1～2年ぐらいは、地域発信においては、デジタルアーカイブを優先的に入れているので、件数はかなり増えて197データベースということになっています。

サイトに行くとどういう197かっていうのが一覧できるようなページがありますので、ぜひ、興味があればそういうところをご確認ください。メタデータはそれほど増えてなくて、大口でないというところで、それでも600万件ぐらい、700万件かな、そのぐらいは増えて、アクセス数も倍以上にはなっているという状況です。もう一桁ぐら行ってもいいかなというふうにつくっている側というか、運営している側は感じていますが、まだちょっとリーチできていない部分が多いかなというふうに感じています。今日こういう機会をいただいたことで、非常にこの辺を盛り上げていくのに役に立つかと期待しておりますので、ぜひ皆様も触ってみてください。

このポンチ絵ですけども、こういう活用户、アーカイブ機関、つなぎ役っていう三者一体で、これがぐるぐる回り出すと、成果物がいろいろな活動、社会的な意義のある、防災に役に立つとか、もちろん教育に役に立つとか、学術研究。今日ここにお集まりの方々は、まずはこの辺が一番のセンターなのかもしれませんが、恐らく社会人で入られている方などは、多分もう一つ地域活性したいよとか、防災に役立てたいとか、深い中長期の目的があってこういうところに、デジタルアーキビストになってみようかなということで入られているんだと思いますけども、そういうところにつなげていくというのが僕たちの目的でもありますので、ぜひ使ったり、ちょっと眺めたりしてみて、こういうところが足りないじゃないかというようなことがあればフィードバックをお願いしたいと思います。

連携と一言で言いますけども、なかなか一筋縄ではいかないですね。190もいろんな組織があって、歴史的な経緯があってためてきたものですから、整理の仕方も無手勝流といいますか、自分たちがこれまで紙の帳簿でやってきたことをそのままエクセルに置き換えて、エクセルですずっとやっているとか、あるいはシステムを導入したとしても金額的な折り合いから非常に単純なことしかできないシステムしか入れられなくて、それに合わせてデータ自身もつくられるということで、もうちょっといろんな情報を入れたいのだけでも、なかなか整理がついていませんというようなところが非常に多いわけですけど、それをまとめ上げる過程で少し底上げするといいますか、割とうまくいっているところに合わせて、この機会に頑張っじゃあメタデータ、あるいは関連情報みたいなものを整備しながら上げていきましょう、あるいはつないだ後でどんどん充実させていこうというようなことにつながることを目指してやっているの、模範役が何かそうでもない人に合わせてデータがどんどん、たまっていつちゃうといいますか、整理された状態が未整理の状態に引きずられていくってことはしたくないですね。

そこで、このつなぎ役っていうのが模範演技の人は模範演技のままつながっていただいて、あとはまだちょっと準備が整っていない人たちも、そこ順次倣っていくような仕組みっていうのをちょっと考えたいわけです。というので、まずはその模範演技ができそうなところ、国立国会図書館自身だとかですね。これは本分野では模範演技をしているので入っていただくとか、文化庁の文化遺産オンライン。これも僕は立ち上げ時から関わっていますけども、これはそれなりに多くの博物館、美術館の情報を入れていると。少なくとも国指定文化財は全部ここで総覧できるようにしてありますので、そういう意味では一つの模範的な文化財の整理の方法と考えられるわけです。

他には、科博などはサイエンスミュージアムネットという科博だけじゃなくて

科学館、科学博物館のネットワークというのがありまして、そこで非常に専門的な標本データの収集みたいなことを整理されておりますので、そういうところとか、こういう模範的なところにまずは立っていただいて、そこがもう既に持っているものを入れていただく。大抵は1館の情報ではなくて、その先に何十館、何百館というものを控えているようなところになっています。

ところが分野によってはこういう図書でもないし文化財、指定文化財とも言えるようなものではないし、科学者が専門的に使う標本などでもないという場合はどこにも引っかけられないわけですね。そうすると、つなぎ役不在ということがほとんどの場合ここになるわけです。

あるいは文化財は文化財なのだけでも、文化庁のこういう整理にちょっとなじまない、学術的なデータベースになじまないってような人は、やはり文化遺産オンライン経由っていうのはちょっと無理で、直接そういう専門的な仲間同士くっついて入れていきたいというようなことがありますので、そういう状況で、下ではそういう個々の状況に合わせてどうやったらうまく、データのこだわりは生かした上で、発見性はそれなりに担保されていくというようなことをしたいということを考えています。

個々のデータのこだわりを生かすと、大抵の場合は細かく分かれて、項目が例えば何とか名称、発掘場所、発掘時期、発掘方法とかって、埋蔵文化財を掘った記録を残すためには非常に必要な項目なわけですけど、そうじゃない掘られたものを整理するには、必ずしもそれは必要じゃないかもしれない。あるいはこの辺の工房でつくられた工芸を整理する枠組みとはちょっと違うわけですね。ですから同じ文化財といっても、大分こだわりが深くなると言葉が通じなくなるということをよく経験するわけですが、ここを何とかうまく勘案しながら集めていくという苦労がありました。現在はこれもいろんなパターンがありますと書いて、それなりに実例もあるのですけども、こういうデータベースを同分野でうまくまとめられる、これはかなり理想形ですが。あるいは同じ組織だけでも、こだわりの違うものをもう既に持っているんで、それをうまく集めていくとかですね。それから同じ分野、テーマの資料を様々な団体・個人から収集してっていったって、ちょっとずつみんなが持っているようなものを収集して、割とボリュームとして意味のあるようなものにしていくというような、収集すること自身がまた一つの活動になるようなつなぎ役の立ち方とか、それからデータベースは持っていないのだけでも、これはかなり、件数として本当は多いわけですね。

我々の連携しているのは200弱ですけども、多分、1000とか2000とか国全体を眺めればいろんなものがあるはずで、研究者が1人いれば昔はデータベースが



サイエンスミュージアムネット

一つできるというようなことを言っていた時代もありますので、たくさんあるわけですが、これを何とかうまく入れていきたい。だけどそういうときにデータベースを個々の施設が持っていない。データだけあるという、データベースって言っても発信するようなデータベースを持っていないってというようなケースもあるので、そういうときには、逆に言うとジャパンサーチ自身がそういう初めてみんなが触ることができる発信のプラットフォームになるような形で、データのやり取りを工夫するというようなことを考えたりもしています。

これは今後状況に合わせて、さらに柔軟な対応を迫られるかもしれません。でもこういうきめ細やかな対応を国会図書館、本当に頭が下がるぐらいなのですが、ども対応していただいている増えてきたということになります。

つなぎ役として、今の何パターンかあるつなぎ役のいろんなパターンとして、こういう31機関ですかね、ここはそんなに増やさないように、先ほど言った契約の件数がこれに縛られますので、ここが何百になっちゃうともうそれだけで事務手続きが大変になってくるので、できるだけ抑えながら参加データベースは増やしていくっていう努力をしているのですが、こういうところは立っていただいて進めていると、多分かなり名のある、もちろんジャパンサーチなどよりも随分前からこういう活動の意義を見だして、実際に進めておられたような組織になっています。

## 2. 活用基盤としてのジャパンサーチ

紙とか実資料を集めるだけじゃなくて、部分的にはカタログの公開も既に行われているような組織というのがここになっています。こういう人たちが味方になってくれたってことが、今回ジャパンサーチ、それなりに成功っていうのもおこがましいですけども、それなりにうまくできて、日本ではほかにライバルがいないようなサービスになったというふうに思っていますが、こういう方々の蓄積を我々にちゃんと出していただいたといいますか、連携していただいたということになります。

ちょっと余談になりますけども、国のこういう委員会っていうのは、大抵は有識者会議とか、僕も有識者であんまりないかもしれないですけど、有識者会議という名前のところがいっぱい出ていますけども、大抵そういうところは言いつ放しで、コメントいただいたということが役人にとっての一つの言い訳になると思いますか。だからこのまま進めますねっていう承認のための一つのステップぐらいにしか考えられてなくて、なかなかそこで意思決定するっていうことは難しいわけですけど、今回は先ほどの実務者検討委員会だったかな、何をやるのか分

からないような名前にしていただいて、そこには本当に実務者だけを集めていたのだですね。だけっていうのはおかしいけど、各組織のうち館長とか副館長が来ても現場のことはほぼ分かっていないので、そういう方々で議論しようっていてもなかなか本当の議論にならないわけです。データを出してくださいって言ってもどうやって出したらいいか分からない人が、じゃあ出しましょうって返事をしても実際に現場は動かないってことが僕らはいろいろ経験していたので、そうじゃなくて本当にデータをグリップして、館の方針などもこういう目的に合わせて改善していく、改定していく意欲もあるような方々に集まっていただいて、その代わり個々の職場が抱えているいろいろな実情、こういうものは博物館の分野ではちょっと無理ですよとか、写真の分野だとこういうところをクレジットしていないと、もうとても一枚も見せられませんとか、そういう個々のドメインの実情とかについてみんなが一つのテーブルで話せるような場所っていうのを目指して。

だから最近はずっとルーチン化しているので、それほど腹を割った本音の議論というのが出ているかどうかちょっと怪しいですけども、最初のほうは何ができるのかちょっと分からないっていう状況だったこともあり、各分野の実情を聞きましょう、あるいは委員じゃない人も呼び込んで毎回毎回そういうもっと深い実情について報告いただいて、それについて皆さんは自分たちの館を顧みながらどうやっていったらいいかっていう相場を考えていくっていうようなことをしました。

そういう枠組みというのが結構重要で、だからこれ、データベースでデジタルアーカイブだから人がたいしていなくてもコンピューター同士がつながればうまくいくでしょうとか、ファイルだけ送っとけばいくでしよっていう話ではやっぱりないんですね。それを支えて、そこにそれなりの情熱をつぎ込んでいる人たちが、じゃあ手を携えましょう。それは面白いかもしれない、自分たちにとってもプラスになるし、間違いなく集まったところでは我々のサービスだけじゃないプラスアルファが得られそうだねと、新しい広場ができそうだっていう期待で皆さんご協力いただいているのだと思います。

というので、その輪を広げていきたいので、ぜひ今日集まった皆様でそういうものがちょっと思いついたりする方は、何でジャパンサーチにこれは入っていないんですか？というようなことを呼びかけていただいて、利用者の立場から、あるいは運営に関わる立場からぜひ後押ししていただければと思います。

これがさっきの数とちょっと違うのですが、ちょっと探したらこういう表になっているのはこれしかなかったんで、ちょっと2年前ぐらいですけど、1年半

ぐらい前ですが、141データベースのときのブレイクダウンを出しています。これ、分野って書いてあるのはそれほど厳密ではないです。国会図書館サーチで持ってきたから書籍関連だねとか、そんなふうになっていますが。書籍っていう一般書籍みたいにして集められているもの。でも、よく見ると青森県立図書館デジタルアーカイブとか。図書館だからいいのかな。この中に県史が入っていたりするのかな。地域資料はちょっと切り出していますね。図書館とか役所が管理しているような本でも、地域資料ということでそれなりの追加情報がついて、データベースでまとめられているものはこういう、青森県史デジタルアーカイブみたいな形で切り出されているとか、こういうデータベースの単位で何かラベルというか分野を規定するとすればこんな感じかなと。これは、美術と人文は完全に別とかそういうわけではなくて、美術、データベースと言われる中に入っているものが、実は書籍でも関係しているとかその逆とか、いろいろあるわけですけども、こんな観点で集められ、つくられたデータベース、デジタルアーカイブが集約されつつあるということになります。

かなり広がりがあるということは見えていただいて分かるかなと。どうしてもこういうデータベースとかデジタルアーカイブっていうと、写真がきれいなものとか、そういうところが自慢をするので、そういうコンテンツで例示してこういう新しいものができましたっていう話をするので、どうしてもビジュアルなものが何枚入っているかとかに引きずられがちで、文化財なら文化財がいっぱい集まっているのだろうと。それ以外の方分野はあんまり関係ないのかなって思われるかもしれないですが、ジャパンサーチに限って言えば、あまり画像が必ずなきゃいけないというようなことは要請していません。

いや、ぜひ入れてもらいたいっていうのは実はあるんですけども、例えばサイエンスミュージアムネットっていう、これは先ほど言った科博を中心とするネットワークですが、ここは世界の標本のネットワーク、標本コレクションのネットワークにも参画しているジャパンのノードなわけですが、そうするとラテン名とか非常に専門的な情報がついている標本の一覧だったりするんですけど、カブトムシだけで何千もエントリーがあるような。ところが写真が一枚も、これ、ジャパンサーチには提供されていないんですね。向こうに飛んでいってもほぼほぼ写真はないんじゃないかと思うのですが。専門家同士は写真が見えるような関係っていうのを築かれているようですけど、この分野ではまだそういうものは一般的でないということで、ラテン名の公開という専門家情報の公開にとどまっているわけです。

こういうのは素人から見ると、その辺で見つけた珍しいカブトムシをこれ

は一体何なんだろうっていうのを判定したりする情報にこの辺が役に立つといいのにといいふうに思いますから、写真の1枚も2枚も入っていたらいいのとは思いますが、なかなかそれは実現していないところです。

それぞれに分野に事情がありますので、仕方がないといえば仕方がないですが、舞台芸術みたいなものも、これはenpakuですかね、早稲田の演劇博物館等から演劇系のものもいっぱい入ってきます。ですから演劇系っていうのは当然まだ著作権が活着ているものが非常に多いです、そういうもののデータっていうのがこういうところでさっと見えるかっていうと、ビジュアルなものが見えるかっていうとそうでもないということです。

ただ、舞台芸術などは今日もお話しされる福井さんなどが中心になって、舞台芸術のデジタルアーカイブ化っていうのはずっと進んでおりますので、福井さんの話に出てくるかどうかはちょっと分からないですが、分野ごとにかなり予算もついて国が応援すればそういうものも進むんだなっていうのを今まさに実証してもらっているところですので、そういうのはいずれここに充実して入ってくるんじゃないかと期待しているところです。福井さんなどとはそういう話もよくしているわけです。

### 3. メタデータの連携方法

次ですが、じゃあいろんなところで集まってくるものをどういうふうに、どの範囲をジャパンサーチはじゃあ集めているのかっていう話をちょっとします。コンテンツ、これが個々の博物館、美術館がデジタル化して持っている実態だとします。高精度の写真だとか、3Dの最近だったらデータだとか、いろんなものを持っているのだと思うのですが、修復のためだけに役に立ちそうな、ピクセルの写真とか、そういうものも入っていると思うのですが、それを全部ジャパンサーチに上げられても大変なことになっちゃう。上げたいっていうところもあるんですけど、僕たちが消さずに国会図書館がそういう写真を未来永劫お預かりしますっていうのはなかなかできそうでできない約束なので、コンテンツについてはちょっと預かることをためらうわけですね。

とはいえ、ビジュアルな、例えば浮世絵っていうのがタイトルだけでサムネイルの一枚もないと、どんな浮世絵か分からないし、果たしてその館まで見に行くと、デジタル的であれ見に行くと、大きい画像を出してくださいって出してもらおう手間といいますか、そこ、この作品を見てみたいっていうふうに思うセレクションのときに役に立たないだろうということで、できたらこのサムネイル、プレビューといいますか、ある程度の大きさのものは出していきたい。

ここでも、あんまりルールを決めるとじゃあそれだけの大きさがあればいいんですねって言って、逆に縮めてそれに合わせてくるっていうようなところが当然あるので、できるだけあまりこっちでサイズ等は指定せずに、できるだけベストエフォートで出してくださいっていうような話をしております。さっき言ったように分野によっては一切出さないというようなところもあるということになります。

メタデータは逆に言うと必須で、これは本でいえば書誌情報のようなものです。タイトル、著者、出版社、出版年みたいな、そういう基本情報はぜひとも入れてください。これは福井さんの話でいずれ出てくるとは思いますけど、メタデータのそういうものは大体著作権がないというふうに認められていて、自由に流通することに適していると。最近サムネイルも一部は入れられるというふうに、著作権を主張できないというふうになっていますけども、大きいサムネイルは少なくともその範疇から外れるので、日本の法律ではサムネイルについてはやっぱりその提供館なり権利者がそれを出しても構わないよっていうことを認めるような仕組み、あるいは権利制限が加わったような状況でしかこれは追加できないというふうになりますが、この青い部分というのをジャパンサーチは集めています。

ですからタイトルとかその他の基本情報で調べることができて、その調べた結果がサムネイル等で一覧できる。場合によってはサムネイルつながりで、この山が書いてあるやつは何だろう？っていったって探しにいくというようなこともできるようになっているということです。この辺をどういうふうに整理したらいいのかっていうのは、やっぱりガイドラインみたいなものをこの委員会としてはつくって、リバイズもここは今ちょっと古い版が上がっていますが、何年かに一回はリバイズしながらこれを提供しているというところですよ。少し専門的な話に興味がある方はこの辺を見ていただければと思います。

もう一つのこだわりは、その利用の区分といいますか権利区分です。特にサムネイル等で、これは多分行った先でどういう形で利用できるものがぶら下がっているのかっていうことも含めての表示ですけども、少なくともサムネイルをどういうふうに扱えるよっていうようなことについて、ここに権利情報を表示するように、できるだけ表示してくださいということにしています。表示がないときは、あるいは該当なしとかって一番厳しめにしちゃっているときは、排除はしていないのでそれでも可能ですけども、できれば個々の作品について個別にできるだけ権利状況が緩いもの、できるだけ活用ができるものっていうのを挙げていた



早稲田大学坪内博士  
記念演劇博物館

だきたいということで、個別に、作品なら作品個別にこれを振ってくださいということをお願いしています。

これは非常に重要なことで、これまでですとデータベース全体でこのデータベースに入っている画像はこういう権利状態でやりますという、だから商用利用は駄目ですとか、これに限っては利用できますっていうことをうたうっていうのが従来のプラクティスだったのですけども、データベース全体で1つルールを考えようすると、当然その中の一番厳しい、まだ著作権が生きているものが例えば1000件のうちの3件あるから、その3件は駄目だから1000件全部駄目にしましょうっていうふうに悪いほうに引きずられちゃうのですね。

悪いっていうのもおかしいけど、厳しいほうに合わせちゃうっていうことが、これは不思議なものでデジタル情報屋さんが悪いのじゃない？と僕は思っているのけども、その権利を決めるなどというのは面倒くさいから、そういうところが一番安全サイドに振りましょうっていうのが大抵の場合の判断だったわけです。

その結果、これまでは美術館のデータベースっていうのはほとんど、画像は個別に、もう著作権が切れている江戸時代のものでも使えませんかっていうふうな形ですね。個別にお問い合わせくださいなどとなっていて、聞けばこれは切れているからいいですよっていう、答えが返っているのですけど、そこの手間を払いながら検索結果を見るっていうのはあり得ないことなので、結局は簡便には見ることができないというところになっていたわけですが、今回このジャパンサーチにせっかくばらばらに集める、個々のものが個々のものとして見える形でデータを集めますので、それぞれに権利の情報を付与してもらおうということをお願いしています。

最初はエイヤツで一番安全サイドに全てを振るかもしれないですが、これはほかの館ではオーケーが出ているのに何でおたくの館は駄目なのですか？というような話でプレッシャーをかけることもできますし、その提供館に気づきを与えるということができるので、少しずつよい方向に動いていくというふうに期待しているわけです。

そういうことでジャパンサーチの2次利用条件っていうのはこんな感じでその条件によって絞り込むこともできるし、個々の作品のページに行くと割と分かりやすい表示がありますので、ああ、じゃあ安心して使いましょうっていうような形で利用が進むと期待されます。

結構時間がたってしまいましたが、活用基盤としてのジャパンサーチ。これ、じゃあ何のためにつくったかという、先ほど言った活用者がどれだけうまく

使って、上に書いてあった社会的な課題の解決につなげていくっていうのが、大上段の目的なわけですが、少なくとも活用者が面白がってもらえるようなものにならないと駄目だろうということで、幾つか、最初はあんまりこういうものは国会図書館が提供するサービスにはなじまないかな、これまでの国会図書館のサービスを知っていれば大体想像がつかますけども、サービスしているものが厳密な意味で正確な情報が提供されていて、いつでも同じ振る舞いをして、淡々といい仕事をするっていうのが国会図書館的なスタイルなわけですね、まあ美意識なわけですけど、ウェブサイトにもそれが反映されていて割とこう、取りつく島がないといえますか、サービスが多かったわけですけど、ジャパンサーチはやっぱり国会図書館が主たる運営者でないということと、これまで国会図書館が付き合っていたのではないコミュニティの人たちとうまく付き合っ、その利用を盛り上げることが重要だから、もう少し柔らかいところに行きましょうよということをさんざん言って、最初はそんなのはどうかしら？とかいうのが多かったのですが、大分今では国会図書館の関係者、担当者は積極的にこういうものを、取り組んでいただいています。

#### 4. ジャパンサーチ API の活用例

活用っていうのは探す、検索するのはまず一番の機能としてもあるでしょう。ただ、この探すのも従来の書誌検索のようにタイトルはこの箱に入れてくださいとか、説明文の一部はここに入れてください、時代はここに入れてくださいみたいなフィールドが分かれていて、それぞれにそれぞれの対応するものを入れないとそれなりにうまくいってこないっていうのはちょっとしんど過ぎるので、少なくともグーグル的な1つの窓にぽんと何でも入れるとそれなりに出てくるとか、それからそういうメタデータで引くだけじゃなくて、画像つながりで画像の近さで引いてくるというような、探すっていうようなこともできたらいいだろうと。

精度はそれほど高くはないかもしれないですけど、気づいていない手がかりを与えられるというようなことでは役に立つかもしれないということです。見つけたものをいきなり仕事に生かすっていうのはありますけど、まずそれを楽しめる、何人かで楽しむとか、面白いのを見つけたよっていうことが共有できるような仕組みというのをつくっていかうということで、楽しむという軸を入れて。

そして「探す・楽しむ・活かす」ということで、だんだん生（活）かす、何らかの割と真面目な役にも立つよっていうようなことにしていこう。これがぐるぐる回るというのを目指そうじゃないかということで、検索機能を面白くすると

か、楽しむっていうのを、模範演技をしてやろうというところで、ギャラリーっていうのを運営者側がいろんな専門家に頼んで、和食っていうテーマで一つギャラリーをつくってくださいとか、刀剣乱舞で刀剣がはやっているから、刀剣っていうのはこんなにいっぱい入っているんですよ、見つかるのですよっていうことで、ちょっとショーケースみたいなものをつくってくださいっていうようなことで、幾つか軸をつくってギャラリーをつくっています。

これは自分たちがギャラリーと同じものをつくることもできるように提供していますので、マイノートっていうので大体同じものをつくっていくっていうことはできますから、あ、こういうふうに集めていけばいいのかと。この、ここに集められている刀剣をネタにもっと文書のほうまで膨らましていこうというようなことで、江戸時代の和書をいっぱい集めていくなどということもできるように、こちらのほうではさらに活用できるものっていうのを提供しています。

1人で使って仲間にちょっと見せるぐらいはマイノートでできるんですけども、一般公開というのはなかなかできないので、グループの中であらかじめバーで共有するとか、あるいはそこに自分たちだったら見てもいいものを、ジャパンナレッジにはまだ入っていないけども、自分たちの研究の仲間では共有できるようなものを追加して、ラボのようなものをつくるというような仕組みも、プロジェクトといいますけど、提供しています。

従って、皆さんが何か課題をやったり、今後どうなのか分からないですが自分たちの周りでそういう活動されるときに、ジャパンサーチに入っているものは一部役に立つようなシチュエーションがありましたら、ぜひこの辺の機能を使ってみていただければと思います。

ちょっと書いてありますけどもこういう感じですね。つくった側のこだわりは、これはカワシマさんっていうNDLの、NDLに最近ハッカーがいるのですが、ハッカーが何人もいるのですが、コンピューターのプロですけど、彼がつくったときのこだわりなのですが、検索窓を普通の窓で普通に引けば、普通に全部のものが引けるのですが、これを裏でチューンすることができて、あらかじめこのデータベースしか見ませんとか、ここに入れられたテキストはフィールドでいえばこの件名と何とかだけにぶち当てて引いてくださいとか、そういう検索のファセットといいますけど、ファセットティングをあらかじめフィルター条件みたいに与えて。

そうするとどんな文字を入れても結構ピンポイントでその分野だけの検索にできるというので、検索エンジンをカスタマイズするというようなことができるような感じになっているのもあります。これも別の方はこだわっているんですけ



マイノート

ど、画像検索ということで、アオキさんという方がこだわっていますけど、面白いものが見つかってくると、思いがけない、模様として似ている。土偶は土偶って大体名前がついているから面白くないですけど、形が似ているとか模様が似ているとか、そういうものが見つかってくると、なるほどと思うわけです。

ギャラリーもいろいろあります。少しこだわりはトリプルアイエフ（I I I F）に対応していて、これは画像公開のためのデファクトのようなものです。デファクトスタンダードのようなものになりつつありますが、大きな画像をネットワーク越しに公開したり、利用したりするというとき、当然大分ネットが太くはなったものの、サーバ側にはかなりの負荷がかかるので、例えば高精細な画像を何十万点も公開していて、みんながそれをちらっと見め決められたメンるだけでも全体をダウンロードしなきゃいけないっていうふうになると、とてつもなくサーバに負担がかかってうまく動かなくなるわけです。

ところがよく考えると高精細な画像を今こういうデジタルな環境で取り寄せて見るときってというのは、いつときにはほんの一部分しか見ないわけですね。その画面の部分しか見えないわけだから。あるいは画面全体に絵全体がフィットするように見たい場合は、どちらかといえば高解像度の絵は必要ないわけです。このフィットするぐらいの解像度があれば十分ということを考えると、今ユーザーの環境が本当にどこをどのぐらいの解像度で求めているのかっていうことのやり取りをして、その必要最小限のものだけを送り込むようにすれば、そんなにサーバにも負荷がかからないし、見ている側もそんなに待つことはないということが起きますので、そこのルールをつくりましょうということです。

サーバを、データを発信する人と、イメージサーバーといいますけど、それから見る側、ビューアーといいますけど、ビューアーの技術と、これがうまくかみ合ったらうまくいくよということで、それを規格化といいますか、国際規格になっているわけではないのですが、I I I Fのコンソーシアムっていうところは世界中の主要なそういう発信者を仲間に入れて、ルールを決めている。みんなのフィードバックによって、それをどんどんリバイズしていくというようなことをやっています。これに準拠しておりますので、この辺の写真はズイズイってやると大きくなるようなものがかなり含まれています。

マイノート。こんなものですが、後でちょっと眺めておいていただければ、ちょっと使ってみようかなっていうふうには。これはちょっと申込みが必要のようですけど、あまりビジネスにすぐ役に立てるみたいな、商用が表に立っているとちょっと許諾を出しにくいですけども、学術的な興味でやられるというような活動であれば止めることはないと思われしますので、ぜひ手続をしてご利用いただけ

ればと思います。

ですので、ジャパンサーチはデータベースやギャラリーっていうのは公開レベルに入っていて、プロジェクトはプライベートな空間で、これは参画、提案して利用しているメンバーだけで共有されていくというような形になります。実はこれ、海外の進んだ図書館などは大抵こういう仕組みを持っています。だから僕はスタンフォード大学の図書館っていうところに入出入りしていますけども、そこなどは研究者がやっぱりプロジェクトをライブラリーに対して提案して、あ、いいですよということになると、ライブラリー側に専門家、その分野の専門家はいっぱいああいう図書館だと雇っているんですけど、サブジェクトライブラリアンと呼ばれる専門家が担当となって、研究者の要望に合わせてどういうデータを集めてきたらいいかっていうことを考えて、こういうカスタムなデジタル利用環境をつくります。

それはクローズドですので、公開だとなかなかできない権利上できないようなものも、クローズドだと相当できる範囲が広がりますから、そういう形で研究を進めるというので。この分野では必須の機能になっているということを聞いております。ですので、国会図書館でも少し試してみようということですね。

## 5. ジャパンサーチが目指すもの

最後にちょっと駆け足になりますが、ジャパンサーチのメタデータ連携と活用。メタデータを集めてきて引けるようにしますっていう話なのですが、これ、提供者のところはさっきもちょっと言ったようにこだわりの項目がいっぱい入っているので、それを全部引けるようにするっていうのはなかなか大変です。かつ統一感もなくなっちゃうわけです。

従ってジャパンサーチでは、大抵こういう類の項目はあるでしょうっていう共通項目ラベルっていうのだけ用意して、そこに、それはフィールドを指定して引ける。後のものは割とどれでも検索には多少利用しますけども、基本はそここのこだわりは引き継がないと。だけど、もともとのところにどういうふうに書いてあったかは調べようと思えば調べられる、見ようと思えば見られるという状況にしています。上の情報を使って検索とかマイノートとか、この辺はつくられます。下の情報、深い情報を使ってAPIを経由して深い情報を出そうと、取ろうと思えば取ることもできるというようなことにしています。

ここはAPIというのはプログラムで呼ぶこのサービスを、別のサービスから裏のシステムとして、裏のデータベースのようにして呼んで利用すると。検索をこっちのプログラムが検索して、結果をこっちのユーザーに返すというようなと

きに使うわけですけど、一ついろんなサービスをつくっていくコンポーネントとしてジャパンサーチを使うこともできるような形で工夫しています。

ここの共通ラベルに絞るっていうのは、次のページにたしかよく書いておきましたが、例えばAっていうデータベースはこんなフィールドになっている。Bっていうデータベースはちょっと取ってあるものが違うので、材質とか技法とか、ちょっとフィールドの名前も違う。これは作品名と言っているけど、本のほうは書名って言っているとか、作者と言っているがこっちは著者と言うとかって、そういうふうに名前の呼び方が違ったりするのですが、これも大体基本的にはこのタイトルとか作成者とかもう少し取ると思いますが、そういうところについては共通だねといって、こういう対応を取るわけですね。

ここの赤い部分だけを利用して、実際のサービスがつくられるという形にしています。マッピングといいますか、こういうふうに全部焼き直しちゃって、元のデータは全部潰してその他に入れるっていうのが、これは従来のマッピングの方法なんですけど、僕たちはこの必要そうなものをコピーして、デュープしてタイトル、作成者という情報もつくります。けども大元の、これは実は作品名でしたとか、これは実は作者でしたっていう情報はそのまま保ち続けるというこだわりが。そうすると例えば映画だったりすると、映画の監督なのかラインディレクターなのか何とかなのかっていう貢献がいろいろ違いますので。原作者なのか。そういうようなものも、こちらの紫のところにはいっぱい書かれていて、それは映画を専門的に見ていくときには非常に重要な情報なので、それを捨てたりはしないということになります。

そういうのをユーザーが自分でこうやって登録するような仕組みをつくってあります。何かちょっと試したい方がおられたら、ぜひ問い合わせしてみてください。誰でも受け付けるっていうわけにはなっていないですけども、どういう形で整理したらいいかの参考の情報っていうのは少なくとも得られますし、それから適切なつなぎ役を紹介していたりして、接続につながっていくという感じです。

さっき言ったプログラマーがジャパンサーチをデータベースのようにして使うことができるよ、そういうAPIっていうのがありますが、この仕組みを使って、これは実は僕が入っている5人、6人ぐらいのフラットなチームでつくったんですけど、カルチュラルジャパンっていうのを試しにつくってみました。ここにジャパンサーチっていうのが入っています。ジャパンサーチに入っているものを整理したやつが、このジャパンサーチRDFストアっていうのに入っています。これは外からたたいたら返事をしてくれるものですが、そういうものがあります。この後、時実さんが多分お話しするヨーロピアーナとかDPLAってい

う、これはヨーロッパを束ねているもの、これは北米を束ねているものですが、そういうものも同じようにAPIを提供しているので、そこからRDFストアっていうのをジャパンサーチと同じ乗りでつくることができます。同じ乗りでこうやってつくると、これはレジストリといいますか、データベースと呼んでもいいのですが、いろんなクエリに答えてくれるデータの入れ物ができます。

これは皆さん知っている方はリンクトオープンデータ（Linked Open Data）とかRDFトリプルとか、そういう形で整理したもので、RDBと同じようなクエリを書いてここから取り出すことができるというような仕組みなんですけども、ここにまず入れます。そうするとこれとこれを合わせたものをデータとする新しいサービスをつくることができます。

これは僕らカルチュラルジャパンと、何かおこがましいですけど、文化的な日本の情報が入っていると。ここはビジュアルにこだわって、絵がきれいなものしか仲間に入れないということを僕たちがプライベートにやっている、プライベートっておかしいね、ボランティアにやっているプロジェクトなのでどういう制約も受けないということで、ちょっと勝手に取捨選択させてもらってサービスをしているというようなものになります。

ちょっとこのつくり方が実は自慢なのでこういう絵を描いてありますが、ヨーロッパの先にぶら下がっている3000館ぐらいものとかDPLAもかなりの館数がありますけど。あとはI I I F対応であれば個々の館がばらばらでも集めてくることはそんなに難しくないで、そういうメトロポリタンとか、モマ（MoMA）とかブリティッシュミュージアム（British Museum）とかそういうところからも集めてここに投入しています。全部投入しちゃうと世界中のものを集めちゃうことになって無理なので、カルチュラルジャパンって切ったのは、ジャパンとかジャパニーズとかジャポネとか現地の言葉で引くと引っかかってくるような項目をここから取捨選択して集めてきて、これをつくったということになります。

これもこの辺をクリックすると実際のサービスに行って、結構面白いことができるという感じです。表示したものを、調べたものをいきなりミュージアムにするなどという仕組みもあって、ここのエクスポートってやってこの箱をクリックすると、今の検索結果が全部展示室に分かれて3Dのミュージアムができるっていう。

皆さん手元で動いているかちょっと分かりませんが、それで3Dでウォークスルーできるという、なかなか活用といえば活用ができるようになっています。検索の結果のあれでもいいですし、自分で拾ったマイボックスっていうような

ところにお気に入りを集めて、それをミュージアム仕立てで眺めるなんていうこともできるようになっています。ちょっと試していただければと思います。

最後、ジャパンサーチが目指すもので、時間になっちゃいましたね。一応、ジャパンサーチはサービスとしていろいろやっていくのだけでも、そもそもどういふところを目指していくんだって。デジタルアーカイブをどう社会に位置づけて、僕たちの生活の中で当たり前のものにしていくのかっていうことをちょっと掲げて、5年間の戦略方針っていうような形で書き物をちょっと出しました。

ここで3つの価値が重要だと。デジタルアーカイブっていうのは、別にジャパンサーチに関わりがなくても、この3つの軸が重要だろう、価値が重要だろうということをちょっとうたいました。記録・記憶を継承していったって、またその整理の仕方を新しく変えていくっていうようなことは当然重要だろうし、それからコミュニティを支える共通知識基盤というのはそこから生まれてくるのじゃないか。これが今のキャンセルカルチャーですか、知らない間にある種の情報が全く見えなくなっているとか、それで利益が誘導されたり、僕たちの認識自身がコントロールされちゃうっていうようなことが起きつつあるわけですけど、デジタルアーカイブに価値があるとすると、もう何十年、百年前の情報が割と包み隠さず今は見えている。

それから今今（いま）の情報も、どうしてこの辺は欠落しているのだろう？ということにユーザーが気づいたりするというようなところまで行けば、こういう知識基盤は非常に重要だと。知識基盤自身がキャンセルカルチャーの総本山になってしまっただけでは意味がないんですけど、そういうところ。こういうのもあって、みんなが草の根でつくること、つくったものが集まって全体ができるっていうことの価値っていうのがそこにあるんですけど。それからそういうものが新たな社会ネットワークの形成になると。個々でそれなりのこだわりで小さくつくっているものが、遠くのもの、同じようにつくられているものとジャパンサーチとかそういうサービスを通じて出会うことによって、お互いがお互いを知ることによって、実は違う立場があるとか、実は同じ立場で別のところで頑張っている人たちがいるとか、そういうところに気づいて新たな人と人のネットワーク、活動と活動のネットワークが生まれていくっていうところが価値なんじゃないかということで、この辺、ちょっと僕らがかなり議論したので熱いのですが、何でデジタルアーカイブなんかやるのだろう？というときには、少しこの辺に帰っていただければと思います。

これで終わるのですが、イサム・ノグチっていう人がこんな言葉を言っていて、ちょっと好きなんですけど。我々っていうのは、我々人間は僕たちが何をこ

れまで見て生きてきたかということの一つのランドスケープにすぎない。僕たち個々の人間は一人ずつが一つの景色だということを言っているっていうのはなかなかいいかなというので、その景色が共通部分を持ったり、お互い同じものをいいねと思っているっていうような別の景色とつながっていったりするっていうのが重要かと思います。では時間が超過しましたが、これで終わります。どうもありがとうございました。

(文責：久世)

## 課題

1. ジャパンサーチについての課題について説明しなさい。
2. ジャパンサーチ API の活用例について具体例を挙げて説明しなさい。