

第 11 講 世界のデジタルアーカイブの発展とその活用

時実 象一（東京大学大学院情報学環・岐阜女子大学）

デジタルアーカイブの種類っていうふうに考えていったときに、書籍、文書、新聞それからテレビ・放送、映画、音楽・音声、舞台芸術、写真、それから美術品があります。その他にも、それとかあとはウェブページ、ゲーム、ソフトウェア、その他というようなのがあります。ここでは、世界のデジタルアーカイブの発展とその活用について考えます。

【学習到達目標】

- ・世界のデジタルアーカイブの動向について説明できる。
- ・世界のデジタルアーカイブを俯瞰して、その活用の変化について具体例を挙げて説明できる。

デジタルアーカイブの種類っていうふうに考えていったときに、ここにあるようないろんなものがあるだろうと。書籍、文書、新聞それからテレビ・放送、映画、音楽・音声、舞台芸術、写真、それから美術品です。この辺が典型的な文化遺産なわけですけども。

それとかあとはウェブページ、ゲーム、ソフトウェア、その他というようなものがあるというふうに考えます。

ここでまずポータルサイトっていう、今日も今、高野先生にお話しいただいたジャパンサーチ、それからカルチュラルジャパン、こういうのはポータルサイトと呼ばれるものですが、世界的に有名なものとしてはヨーロッパのナショナルデジタルアーカイブ（NDL）とDPLAがあると。あるいは国立国会図書館でやっていたいっている東日本大震災アーカイブというのがあると。

それと違ってデジタルアーカイブサイトっていうのは、これは実際に物を集めているところです。ここではインターネットアーカイブとウィキペディアについてちょっとご紹介したいと思います。

1. 書籍

本ですけども、本のデジタル化っていうのは国立国会図書館さんが随分やっていて、今日はちょっと世界ということで国立国会図書館の話はちょっと割愛



国立国会図書館させて

いただいているのですけども、古いところでインターネットアーカイブというところが、2002年、今から20年前にミリオンブックプロジェクトというのを始めて、そこがこの本のデジタル化ってということでは最初なのかなというふうに思います。

これはインドにこういう機械を切って、インドでやってもらったと。当時インドは正直言ってお給料が安かったものですから、インドでやるのがいいだろうというようなことで始めたというような話があります。これは自動的に本をめくるKirtasという装置なんですけれども、これと似たようなものは今でもありますけれども、結構いいお値段がして2000万円とか、こういうお値段がするとなかなか普及はしなかった。

それからこれはインターネットアーカイブって今でも基本的に使っている手めくりの、ですからちょっと分かるかと思いますがV字型になっておりまして、上から本をガラス板で本を押さえて写真を撮るんですね。スキャナーって言っていますが、実際は写真を撮影します。ですからカメラの解像度っていう関係もあります。当時は300dpiが限度だったんですけど、今はもっとカメラが良くなっていますからもっときれいなものが撮れます。

こういうので図書館と今は協力しております。これもちょっと2～3年前の数字なのでちょっと今は違っていますが、特にボストンの公共図書館とはかなり一生懸命やっております。最近、中国とも協力して浙江大学で、やっている。

これはボストン図書館にある、スキャンをやっている様子。御覧のようにテントみたいなのを作って、その中で撮影をしているというものです。この集めた本をどうするかっていうと、実はオープンライブラリーっていう電子図書館をやっております。この電子図書館は誰でも読めると。著作権が切れている本は誰でも読めると。これは国立国会図書館でもそうですけど、生きている本については、所蔵図書館のメンバー、図書館っていうのはみんな図書館のメンバーになれば本は読めるわけですね。

それは日本でも同じですけども、同時に1名だけ読めるということなんですけれども、これは勝手に電子化しているものですから著作権侵害っていうことを訴えられておりまして、現在係争中ということです。これはオープンライブラリーの画像です。

日本で結構有名になったのはグーグルブックスっていう、ちょっと若い方は御存じないぐらいの話になっちゃったんですけど、グーグルがブックサーチという

のを始めまして、これは図書館から本を借りて電子化していると、この進行状況ってというのは本当に何冊持っているかっていうのは、データがありませんで、2008年というもう大分前の数字になっているんですけども、その当時で700万冊あると、恐らく今は、1500万冊以上は電子化しているだろうと思っているんですけども、グーグルってというのはそもそもある意味じゃポータルですよ。自分でデータを持ってなくて、ほかの人のデータを検索できるっていう検索エンジンになるわけですけども、自分でもデータを持とうということで始めたプロジェクトなんです。

これが2004年に始まってもう20年前になっちゃうのですけども、こういうアメリカとイギリスの主要な大学と提携して始めたということです。これも2005年に訴訟が起きていろいろあったわけです。

それでこれは最終的にはアメリカの裁判所で、要するにスキャンしてデジタル化して、その一部分だけ、本当に、スニペットといいますけれども、短い文章を検索のために提供するっていうことは合法であると、フェアユースであるということで決着をしまして、これは今でもやっている。ただ、デジタル化のほうはいろんな事情でもう追加のデジタル化はやらないということになっております。

こういうデジタル化して検索に提供するのはいいっていうのは、日本の著作権制度も、数年、また福井先生からお話があると思いますけど、日本の著作権制度もそうになっておりまして、皆さん、デジタル化するのは何でも許諾を得なきゃいけないと思っていらっしゃるかもしれないけどそれは間違いで、デジタル化して検索に提供するのは勝手にやっていいんですよ。許諾を得ないでですね。アメリカでもそういう判決が出ているということです。

このグーグルブックスでデジタル化した書籍データをHathi Trustっていうところで管理しております。共同管理しています。CADAL（カダル）っていうのは先ほど言いました浙江大学でやっているプロジェクトで、杭州市で、ここでデジタル化を進めている。これは先ほど言いましたインターネットアーカイブが支援しているプロジェクトです。

2. 写 真

今、本のお話をしましたけど、写真のお話をしたいと思います。写真につきましてはどこのデジタルアーカイブに行っても写真はやたらいっぱいあるわけですから特に珍しいものではないのですけど、ここでは商業的に成功しているっていう上でGetty Images（Getty Images）っていうのをご紹介します。

これは皆さんテレビなんかを見ていると、テレビの、例えば有名人が死にましたと。すると写真のところ右下とか左下のところにゲッティイメージって書いてあるのですよ。気をつけてテレビを見ていてください。そうするとこれはゲッティイメージから写真を借りて使っているんですよね。これはもちろん有料で借りているわけですけど、ここで使う分には合法ですから、ものすごくたくさんの写真がありますので、マスコミはもうゲッティ様、ゲッティ様で、非常に多用されております。

これはゲッティのホームページですかね。いろんなスポーツの写真とか何でもあるというわけです。

あとは写真で有名なのはウィキメディア・コモンズっていうのがあって、これはウィキペディアに基本は掲載された画像を収録しているっていう考え方なのですけれども、必ずしもウィキペディアに載ってなくてもいいんですけれども、これのすごくいいところは、撮影者の情報と権利情報が明記してあって、これが自由に使えるというのが基本です。

ですから一番下書いてありますデジタルアーカイブのバックアップとしても利用されているということで、日本でもいろんな方が使っています。これは岐阜女子大学の写真で、このウィキメディア・コモンズに載っている写真ですけど、右下に「CC BY-SA 3.0」って見えますかね。書いてあって、それで要するにCC BYですから誰でも使っていってということで、これはMonamiさんという方が写真撮影した写真ということで、皆さん自由に使えるということであります。ここに書いてありますね。こんな感じで非常に安心して使える画像がありますので、便利なものです。

3. 新 聞

新聞の話をしたいと思います。新聞につきましては、日本については朝日新聞、毎日新聞、読売新聞、あるいは日経さんと皆さんアーカイブを持っていられるのですけれども、アメリカではもっと古いところ、17世紀から新聞をアーカイブしています。

アメリカという国は新聞ってすごく大事で、町ができると新聞ができるわけですよ。それで映画でなんか時々出ますけれども、自分たちの町の新聞っていうのをみんな持っていました。

持っていましたっていうのは、今、それがだんだん潰れちゃっていて、ちょっとこれは問題なのですけれども、そういうことで。その記事に何が載っているかという、もちろん町の出来事が書いてあるんですけど、冠婚葬祭が載っているの

ですね。だから誰それが結婚しました、誰それが亡くなりましたっていうのが必ず載っているんで、これが自分のルーツ探しに使えるということで、大変これ自身が商売になっているということですね。ここにありますように、これは比較的新しい1942年の新聞ですけど、こういうふうに掲載されている名前で検索できるというようになっております。

イギリスでも新聞のアーカイブがありますけれども、イギリスの場合はこういう草の根の新聞もありますけれど、どちらかというとやっぱり重要なブリティッシュニューズペーパーなんかですと、主要な新聞って、イギリスも随分古いですから、ここあるように1600年からのアーカイブがあります。

こういうところで、例えばこれはオックスフォードガゼットっていう新聞ですけども、これはちょっと時間がないのでちゃんと読めないですけどいろんな記事が載っております、例えばこういうふうにはこれは同じ新聞かな、切り裂きジャックっていうジャックザリッパーという有名な殺人鬼がいて、捕まらなかったわけですけども、この記事が載っております。当然ながら新聞に初めて出たときは切り裂きジャックっていう名前じゃなくて、ホワイトチャペルマーズ、ホワイトチャペルっていうところで殺されたのでそういう名前がつきました。

4. 映画

映画の話。映画っていうのは古いものっていうのは全部テープ、フィルムで撮影していたわけですね。これは幾つかのフィルムの種類があるんですけども、物理的にやっぱり劣化するんですね。こういうふうにくしゃくしゃになって縮んじゃっていると。

ここまできちゃうとちょっともう見るのが困難ですね。見ることもできたものにつましましては、こういうスキャナーっていうのがあって、これでデジタル化できます。

これはみんぱく（国立民族学博物館）にある機械の写真を撮らせていただいたんですけど。ただ、撮っただけではなかなか汚いということで、画像修復とかカラーコレクション、こういうことをやるのが望ましい。これは例えばデジタル修復版っていうのはテレビで時々やっておりますけども、“男はつらいよ”の下と上とを見ていただいて、ほぼ同じシーンなんですけれども、全然色の鮮明さが違うし、それから傷の修復はちょっとないですけど、これは色ですね。どちらかというと色の修復をやっていると。顔の色なんかも全然違いますよね。こういうことが、やっぱりやらないとなかなか商業用のものにはできないので、大変手間がか

かるということです。

こういう007とか、こういうもうかるものは皆さんやられますけど、なかなかもうからないものをやるのは大変ですよ、お金がかかるから。フィルムアーカイブってということについては、ヨーロッパのフィルムアーカイブをちょっといろいろ聞いたり、調査に行ったりしたのですが、これはドイツのフィルムアーカイブで、これを見るとどこにもフィルムアーカイブって書いていないですけども、こういう倉庫が後ろにあって、前は家具屋さんなのですよ。家具さんの裏の倉庫を使っているとか、そういう。これはフランスのシネマテークフランセーズっていうのですが、これ、は要塞の、フランスが第1次世界大戦の頃にいっぱい要塞を造っているんですよ。そういう要塞の地下っていうのは恒温恒湿、光が安定しているのでそういうところにフィルムを保存しているとかですね。こういうのが、ちょっとここは飛ばしますけど。

それでオーストリアのフィルムアーカイブというのがあるのですが、ここはこういう地域のフィルムを集めて、それでデジタル化するというプロジェクトをやっております、これは日本でも紹介されたので御覧になった方もいると思いますけど、ちょっと詳細は割愛しますが、興味のある方はオーストリアのフィルムアーカイブということで検索していただくと文献、常石（史子）さんという方が書かれている文献があります。

5. テレビ

テレビなのですが、テレビのデジタルアーカイブとして割と有名なのは、アンダースタンディングナインイレブン（Understanding 9/11）っていうことで、ナインイレブンって言っても、これも若い方は御存じないと思うのですが、2001年の9月11日にアメリカの世界貿易センタービルに飛行機が突っ込んだ。そのときの世界中のテレビを全部こうやって集めてデジタル化して、ここに保存しております。

ですからこんなフィルムですね。これはNHKのそのときのニュースがここに保存されているということで、飛行機がちょうど突っ込んだところの画像が、実際のこういうニュースの様子も見るができます。その続きとして、インターネットアーカイブでは、このTVニュースアーカイブっていうのを今やっております、これは2012年から、このナインイレブンからもう10年たっていますけども、主要なテレビのニュース番組をアーカイブしている。これはアメリカの著作権法で、アーカイブすることが、図書館なんかがアーカイブすることは合法ですので、これで全部保存している。



Understanding
9/11

それからもう一つアメリカのテレビ番組については、各番組には聴覚障害者のための字幕情報をつけるということは、これも法律で決まっているので、それも一緒に入手して表示しているということでもあります。

そうしますと、これは例えばBBCニュースでジャパンランドスライド (Japan Landslide) . これは熱海の実話、あれは2年前かな、3年前かな、熱海の土石流のニュースで、BBCでこういうものが保存されておりますので、もちろん日本に関するテレビニュースはいっぱいありますので、大変調べると役に立ちます。

6. 新型コロナウイルス感染症 (COVID-19) 下の社会デジタルアーカイブ

最近、これはちょっと特殊なものなのですが、新型コロナウイルス感染症っていうのはそろそろ収束（終息）に向かっているというふうに言われていますけれども、これでデジタルアーカイブというのがあると。

これは実際ネットに載っているのを見ると、アメリカが一番多いですけども、この辺がどういうお国柄なのかっていうところまでちょっと分析できていないのですけれども、かなりいろんなものがある。

機関別に見ると、アメリカの場合は歴史協会っていうのがありまして、ヒストリーソサエティーみたいなのが各地域にあって、そういう地域がやっぱり新型コロナで皆さんロックダウンということで家から出られなくなってしまった。

そういう非常に特殊な状況ってやっぱり記録しなきゃいけないっていうことで、そういうことが始まったということです。いろんなものがありますけど、写真だったり、こういう自宅でこれはトランポリンをやっているんだか、トランポリンっぽいようなことをやっている写真とか、それとかマスクの写真とか、こういうもの。

あるいは体験記、実際に地域の方に書いていただいたこういうものは全部保存されております。それから家でこもってやることがないから、こんな絵を描きましたとか、そういうのも保存されているところがあります。これはなかなかやっぱり新型コロナという特殊な条件というものをデジタルアーカイブしようという試みですね。

7. ウェブサイト

それから非常に重要なものというのはウェブサイトで、ウェブサイトについては皆さん御存じだと思うのですが、ウェイバックマシン (Wayback Machine) というのがありまして、ウェイバックマシンというのは、これはもともとこの空

飛ぶロッキーっていう昔テレビ漫画があって、そこに出てくるこういうタイムマシンのことなのですね。その名前を取っております。

インターネットっていうのはどのくらいで始まったかっていうことなのですが、大体1995年頃から商業利用が始まったということで、岐阜女子大学さんでも1998年頃からウェブサイトをつくっていると。インターネットを開始したって、ちょっと言葉はよくないですね。ウェブサイトを開始したと。これは岐阜女子大っていう上のほうに、こことここに、カーソルは見えるのかな、URLが入っていると思うのですがgijodai.ac.jpで、これで検索しますとこういうふうに1998年からずっと保存されているということが分かります。これは保存されているもののうち、1998年12月12日のサイトと、こんな感じで保存されています。

あるいはこれが、これも同じ日かな。リンクで、あ、違いますね。1999年1月25日ですね。これでこういうふうに保存されています。これは世界中のウェブ 사이트がこういうふうに保存されているので、もう非常に貴重なリソースだといえることができます。

8. ゲーム

ゲーム。ゲームっていうのはなかなか保存が難しいのですが、これはやっぱりインターネットアーカイブというところでアーケードゲームの保存をやっています。アーケードゲームっていうのも、これも今は死後になっちゃっているわけですが、いわゆるゲームセンターというのが街角にあって、そういうところでゲーム機が置いてあって、そのゲーム機で流れていたゲームです。100円を入れると何分間かゲームができると。こういうもので、これは有名なインベーダーゲームですが、こういうものが実際に動く状態で保存されています。

以上がいろんな、このほかにもいろんなものがあって、さっきの舞台芸術とかいろんなものがありますけども、時間の関係でちょっと飛ばしたいと思います。

9. ネットワークポータル

(1) ヨーロピアーナ

それでネットワークポータルということについてお話ししたいと思いますけれども、先ほど高野先生からも何回も話がありましたけど、ヨーロピアーナというのがあります。これは欧州のデジタルアーカイブコレクション、例えば高野先生のカルチュラルジャパンの話でも出てきましたけども、欧州、アメリカも日本の



Wayback Machine



ヨーロピアーナ

文化資産っていうのをかなりいっぱい持っていらっしゃるんですね。博物館とか美術館で持っている。そういうものもデジタル化されているものがありますので、こんな感じで、これはUtamaroって書いてありますけど、歌麿の絵と、こういうものがデジタル化されて見ることができます。ここ、日本語が書いてありますのは、これはどこだったかな、National Library of Franceって書いてあります。フランスの国立図書館ですけども、ここでは日本語のタイトルをつけていると。日本人もいるんでしょうかね。でも全部が全部ついているわけではなくて、ローマ字とか英語になっちゃっているものもあるというわけです。

これはどういうものかっていうと、そもそもデジタルアーカイブのネットワークで、ジャパンサーチのモデルとなったということで、そもそも欧州委員会というところが運営しております。欧州各国のコンテンツプロバイダーを結合。ヨーロッパの場合、もともと各国にコンテンツプロバイダーって、要するにデジタルアーカイブを提供する機関っていうのがある程度はあったわけです。

ただ、そういうところもまだ本格的にやっていない。これがどんどん統合しましょうということだと思います。アグリゲーターっていうのはジャパンサーチではつなぎ役って言っていますけども、図書館とか美術館、博物館、文書館、こういうところがつなぎ役になってヨーロッパアーナをつくると。この後で出てきますけど、アメリカのデジタル公共図書館、Digital Public Library of America、D P L A、こういうものとも連携しているということです。

このデータ提供者の分布っていうのを見てみますと、国別のアグリゲーターっていうもののデータ数、これはメタデータ数になっております。ですから件数ですけど1200万と。それから館種別のアグリゲーターっていうのが900万。館種別っていうのは、例えばオランダの国立美術館、ライクスミュージアムって、ここなんかは自分で直接ヨーロッパアーナにデータを提携しているのでこういうのが入っています。

それからこのヨーロッパの場合はE Uプロジェクトというのがあって、これは特徴的なのですが、EUで、ECでお金を出しているわけですね。初めからデジタル化をやらないうってということで、このヨーロッパアーナを通じてお金を出している。

だからヨーロッパアーナっていうのはある意味で、デジタルアーカイブの助成機関になっているんですね。年間20億近いお金を持っていて、これをいろんなところにばらまくなって言葉は悪いですが、配分して、それでデジタル化をやらせているという、こういうことがありますので、EUプロジェクトの件数が結構大きいというようなことであります。

データの種類としてはやっぱりイメージ、写真ですね。写真が圧倒的に多くて、それから文字、これは本です。新聞もあります。本とか新聞です。それからあとサウンド、ビデオ。これはちょっとデータが2015年で古いのですけれども、なかなかこの辺の新しい数字が手に入らなくてちょっと古い数字になっております。

国別ですが、これも2015年でちょっと古くて申し訳ないですけども、フランス、ドイツ、スウェーデン、イタリアというのがあって、この大体EUを引っ張っているのはやっぱりフランスとドイツですからね。フランスとドイツが大変多いということです。

意外と少ないのはユナイテッドキングダムっていう真ん中の右のほうにありますけど、イギリスっていうのは今回EUから離脱しちゃいましたけども、なかなかやっぱり独自路線ということで、やや貢献が少ないということでもあります。やっぱりヨーロッパの特徴っていうのは、ヨーロッパのEUに加盟している、あるいは加盟していない国も含めて、ヨーロッパの非常に多くの国を網羅しておりますので、そこがやっぱり非常にダイバーシティーというか多様性を担保しているっていうことが一つはすごいかなと思っています。

DPLAですけど、アメリカのデジタルアーカイブコレクションで、ちょっと数字が隠れちゃってて申し訳ないですけど、その下に2200万と書いてありますが、これもちょっともう古くなった数字なんですけれども、ごめんなさい、ヨーロッパもDPLAもそうですけど、こういう数字ってなかなか難しく、例えば時々ごみ掃除をやるんですよね。本当にどうでもいいっていったら言葉は悪いけど、非常に汚いデータなんかも紛れ込んでいますので、こういうのを捨てたりすると時々減ったりするんですよね。だからこういう数字っていうのはあくまで目安として見ていただいて、また先ほどヨーロッパっていうのは4000万ぐらい、こっちは2000万ぐらいってちょっと少ないということが言えます。これもアメリカも先ほど言いましたように、大変日本の文化資産っていうのはありますし、これもたまたま歌麿ですけどもこういうものを見ることが出来ます。

(2) 米国デジタル公共図書館 (Digital Public Library of America)

これはDigital Public Library of Americaということでヨーロッパが非常に成功しているので、ぜひアメリカでもやりたいと。ただ、そのアメリカというのは特定の機関を除いて国っていうのがこういう文化事業に関与するっていう伝統がないのです。



Digital Public
Library of America

確かに公文書館とか議会図書館とかありますけれども、例えば国立美術館っていうのはないんじゃないかな。あるかもしれないけど、そういうところはやっぱり違うんですよね。だからどうしたかっていうと、図書館なんか話が合って、私たちでつくりましょうという、ある意味では草の根でつくったというところが大変違います。

したがって、これが2013年4月に始まって、2019年4月にはこういうふうアメリカ中に広がったと。まだ一部参加していない国がある。これも2019年の数字ですけども、この後コロナになっちゃって、DPLAの総会っていうのがちょっとなくなっちゃったものですからその後の数字がちょっと手に入らないのですけれども、大体今はこんな状態になっています。

何を言いたかったかっていうと、ここで図書館の連合体ですから、基本的に国の支援っていうのはないわけです。図書館なんかは、会費を出しても知れているので、これは基本的にいろんなところの補助金を取って運営しています。

ですからちょっとヨーロッパと比べて財政的にはかなり苦しいっていうか、どうやって維持していくかっていう問題がある。もちろんスタッフもそんなに多くないというのですけども、一応かなりこう、システムとして確立している。

そういう意味では日本のモデル、ジャパンサーチっていうのは、これは完全に日本の国のお金でやっていますから大変ある意味じゃ安定している、財政的に安定しているってことでヨーロッパに近いというようなモデルになっています。ちょっと特筆したのはウィキペディアとの連携ということで、ウィキペDPLAっていうのですけど、これはどういうのかというと、一つはDPLA、横浜っていうこれはウィキペディアの英語版ですけど、検索してみますとこういうふうにDPLAのリンクが出てくると。こういうような機能があります。

これはクロームの、グーグルクロームのアドオンみたいな形で実現できます。それからもう一つは、これをやりますと、ここをクリックしますと、この場合はニューヨークパブリックライブラリーのデジタルコレクションから横浜の横浜吉田町より馬車道を臨むという、こういう色つき写真ですね。カラー写真ではなくて後から色をつけた絵はがきですけど、こういうものを見ることができます。

(3) ジャパンサーチ

ジャパンサーチについてはもう既に詳しく説明がありましたのでこれは飛ばしますけど、連携機関とここにありますように、こういうようなところが現在協力



ジャパンサーチ

関係を結んでいるということがホームページに掲載されていますので、見ていただければよろしいかと思います。カルチュラルジャパンについても先ほどご紹介ありましたので先に行きたいと思います。

(4) 東日本大震災アーカイブ (ひなぎく)

特にここでご紹介しておきたいのはポータルとしてテーマ別のポータルっていう、必ずしも多くないですけど、東日本大震災アーカイブ (ひなぎく) というのがあって、これは2011年3月11日ですか、ここに東日本大震災っていうのが発生したわけですけども、これをやっぱりデジタルアーカイブとして保存しようということ、いろいろ議論があって最終的に国立国会図書館が引き受けて総務省のプロジェクトということで始まったんですけど、国立国会図書館が引き受けて運営をしております。

ここにありますようにユーチューブとかそういうところにも全部リンクされていて、これを見てもちょっと分かりにくいと思うんですけど、車が津波に流されたところですね。それを被災者の方が建物の屋上から何かからスマホで撮っているわけですね。そういうものが記録されているというわけでありまして、ここにありますように、動画とかいろんな資料とか、それから写真とかが載っているというわけです。

(5) インターネットアーカイブ

ここでちょっと、これが少し実際にアーカイブをやっている機関の紹介ですけど、インターネットアーカイブにつきましては、さっきから何回も出ていますけども、もともとこれはアメリカのサンフランシスコにありまして、プレシディオというところに最初あったのですよね。

これは御覧のとおりゴールデンゲートブリッジが見えるすごく風光明媚な場所に、こんな感じですけども、ここにこういう建物があって、これがインターネットアーカイブの本部だったわけです。ここですとかなり小さいですから、20人、30人ぐらいいたらもう満杯というようなところでやっていた。

ここにあるのはただで場所を借りられたのでここに落ち着いたというわけでありまして、現在はここから車で20～30分でそんなに遠くないんですけど同じサンフランシスコの中でここが本部になっております。これはもともとアメリカのどちらかという新興宗教の教会だったところを買って、それで本部にしています。こんな感じで皆さん仕事をしていると、これはどういうところかという、まず非営利団体で、ブリュースター・ケールっていう人が創設した。



東日本大震災アーカイブ (ひなぎく)

名前を読むとケールって読みたくなるんですけどケールというふうと呼んでおります。1996年にウェブサイトの収集を開始したということで、細かい経緯は書いてありませんけど、最近「カナダに支部(?)を」ってクエスチョンしているのはちょっと法的にはカナダに独立したインターネットアーカイブのブランチ機関っていうのをつくったわけですね。そこになぜつくったかっていうと、実はトランプが大統領になってもう6年になりますかね。6年前になったときに、もしかして言論統制が起きるかもしれないっていうことを非常にブリュースター・ケールが危惧しまして、やっぱり避難、バックアップをつくらうと。で、カナダにつくらうということで支部をつくるということ。こちらにはサンフランシスコにある本部のデータをコピーして置いておくというようなことをやっております。

ここでやっている主なプロジェクトとしては、さっき言いましたウェイバックマシン、それからTVアーカイブ、それから書籍。この辺はもう既にご紹介いたしました。そのほか動画とか音声と画像、ソフトウェア。最後のソフトウェアも紹介されております。こういう多方面のものを、実際にコンテンツを集めて実際に保存しているというところが、これは博物館と一緒に、自分たちではインターネットアーカイブというのは図書館だというふうに言っているわけですが、そういう機関であります。

(6) フィジカルアーカイブ

ちょっと一つだけご紹介しておきたいのは、フィジカルアーカイブというのがあって、これは本の電子化ってやっているわけですが、この電子化した本っていうのはどうするかということです。

例えば先ほど紹介しましたアメリカの……、あれは細かく紹介してなくて、アメリカの議会図書館ですね。ライブラリー・オブ・コングレスでも新聞の電子化プロジェクトっていうのをやっているのですが、ここでは実はもともとの新聞っていうのはもうないんですよ。それでマイクロ化しちゃって、マイクロフィルムにしたものがあると、それをデジタル化する。そうすると元のマイクロフィルム、あれはもともとの紙の新聞はどうするかっていうと大体捨てちゃうんですね。

どこの図書館も、日本に限らず大変場所がないので、デジタル化するっていうのは書庫を整理するためにデジタル化するんだっていう考えもあるわけです。ちょっとそれはいかがなものかっていうのがこのフィジカルアーカイブの考えで、そういうものを集めて、もうとにかく捨てないで倉庫に保管しましょうと。

これはすぐに見るわけじゃないので、取りあえずとにかく保管しましょうということで、このフィジカルアーカイブというのを始める。

これはフィジカルアーカイブが始まったときのお披露目のときの写真ですけど、真ん中に立っている茶色いジャケットを着ているのはブリュースター・ケールです。後ろにあるすごいコンテナ、これが本を入れたコンテナなわけです。こんなコンテナで、この人が立っているのサイズは分かりますけど、コンテナの中にこういうふうに段ボールをただ積んでいるわけですね。

この中に本が詰まっています。これ、めったに出すことはないわけですけど、取りあえず取っておこうと。こういうふうにラップをかけて、それから一応空調も入っておりますので恒温恒湿にして、それから虫が入らないようにして、それで保存しておく。各本については全部カタログしていますので、何かの拍子で例えばデジタル化したのにこのページ全部飛んでいるじゃないかと、これは非常に貴重な本でぜひ読みたいということがあれば、ここから取り出してもう一回デジタル化すると。あるいはそういうことができるわけで、とにかく捨てないで、そういうのをやめましょうということでやっているわけです。

ブリュースター・ケールは、これは先ほどのプレシディオのところでお伝えしているんですけど、こういう人ですけれども、これはどういう人かというと、もともと人工知能といいますか知識データベースの開発者で、そこでいろいろやったものをアマゾンに売ったりして大分お金を稼いだんですね。

そのお金をアメリカでは起業者っていうのはものすごくお金を稼ぐ、何億、何十億って稼げるわけですけども、そのお金を自分でリタイアして何かカリブ海か何か行ってやるっていうんじゃないくて、この人はインターネットアーカイブを始めた。それで自分のお金で始めたというようなことであります。

(7) ウィキペディア

最後にウィキペディアの話をしたと思います。ウィキペディアっていうのは皆さん誰でも御存じだと思いますけど、御覧のようにフリーエンサイクロペディア、フリーというのは無料だということです。お金の取られないエンサイクロペディアということで、英語が620万件、これは1年ぐらい前の数字だと思うんですけど、日本語が120万件とこういうような数で、多いところかというと日本は必ずしも上じゃないですけどトップ10には入っているのかなというふうに思います。

これはジミー・ウェルズっていう人が2001年に始めたものですけども、ですからこれはもう20年以上ですね。もともとインターネット百科事典というこ



ウィキペディア

とで、百科事典だから内容をやっぱり詳しくちゃんとチェックして載せるべきだっていうふうに誰でもそう思うわけです。だけど、そうすると記事が集まらないっていうか、やっぱりチェックに時間がかかるわけですね。当たり前のことで、それをもうやめようと。要するにとにかく書いたものは全部公開して、それをみんなが訂正するような仕組みをつくった。

これがやっぱりインターネットの、一応クラウドという考えですけども、誰でも訂正できるっていうところが大きかったですね。それで急激に記事が増加して現在に至っているということです。この絵はジミー・ウェルズですけども、先ほども言いましたようにウィキメディア・コモンズというのが姉妹プロジェクトとして、これはデジタルアーカイブの保管庫になっているということで、これはもう紹介いたしました。

(8) ウィキペディアタウン

ここでウィキペディアタウンについてちょっとご紹介したいと思いますけど、これは世界的なプロジェクトですけど、日本では結構盛んに行われています。これは地域の有志が集まって地域に関するウィキペディアの項目をつくる、あるいは項目を編集するということでイベントとして実施していると。これで集めた写真とかそういうデータをウィキメディア・コモンズに搭載して、これをギャラリーとして閲覧できるような仕組みになっています。

例えばこれは岐阜県岐阜市弥八町にある地蔵の弥八地蔵っていうんですけど、私がちょっと現物を見たことがないのでですけども、こういうふうにページがある。このページはこのウィキペディアタウンっていうのでページを、ゼロからつくったのではないと思うのですけれども、編集している。ここに一番上にギャラリーがありますね。

こういう写真は新しく撮ったりして、それで編集すると。こういうことをやっているわけですね。これがその写真の一つですけども、ここに、これはウィキメディア・コモンズに載っている写真ですので、右下に「CC 表示・継承 4.0」と書いてありますけれども、荒川宏さんという方が撮影して投稿しているわけです。

ですからこれはCC BYですので、自由に皆さん使うことができます。こういうふういろんなウィキペディアタウンがあって、これは函館市中央図書館のウィキペディアタウン、これは和束町っていうのですけど、どこだったかな。忘れちゃいましたけど、そういうのがある。ちょっと調べてみたんですけども、このコロナの中でこういう人が集まるというのはなかなか難しいんじゃないかと

思ったら、去年だけでもすごくたくさんやったんですね。

ざっと50～60は、2022年の1年間だけれども日本中でこういうのが行われています。

そうすることでデジタルアーカイブというのはいろんな種類がありますよと、デジタルアーカイブのポータルサイトとか、その他のデジタルアーカイブにはこういうものがありますということをご紹介いたしました。

(文責：久世)

課題

1. 世界のデジタルアーカイブに関する課題について説明しなさい。
2. 世界のデジタルアーカイブを見て、コングデジタルアーカイブする対象として何があるか具体例を挙げて説明しなさい。