

Title	アーカイヴィングにおけるコンテキストの共有
Sub Title	
Author	金子, 晋丈(Kaneko, Kunitake)
Publisher	慶應義塾大学デジタルメディア・コンテンツ統合研究センター
Publication year	2014
Jtitle	慶應義塾大学DMC紀要 (DMC Review Keio University). Vol.1, No.1 (2014. 3) ,p.18- 22
JaLC DOI	
Abstract	
Notes	特集：DMC研究センターシンポジウム：第3回 デジタル知の文化的普及と深化に向けて： コンテンツとコンテキストの統合的アーカイヴィングに向けて
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO32002001-00000001-0018

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

アーカイヴィングにおけるコンテクストの共有

金子 晋丈 (DMC 研究センター研究員 理工学部専任講師)

慶應義塾大学の金子でございます。本日は「アーカイヴィングにおけるコンテクストの共有」ということでお話をさせていただきます。まずはじめビデオをご覧ください。

ビデオ音声：一般的なコンピュータのデスクトップ環境ではファイルはフォルダに分類、保存され、ツリー型の階層構造となります。何か必要なファイルを探すときには関係があるファイルやフォルダを見つけるために上の階層に戻り、他のフォルダから探します。または適切なキーワードで検索しなければなりません。関連するファイルが複数のフォルダにバラバラに保存されている可能性もあります。このようにファイル間に存在する関係を知ることやあるファイルが他にどのようなフォルダに配置されているか知ることは困難です。

本研究 MOSAIC II システムはファイル間の関係を可視化します。ファイル間の関係は自然言語を用いず、有効グラフのみを使用して表現され、カタログとしてデータベースに保存されます。複数のユーザがそれぞれの視点から異なる関係を記述すればカタログのデータベースを介してそれらを共有することができます。これによりファイル間の関係の解釈にユーザごとに異なる広がりを持たせ、発展性のある多様な視点の共有が実現可能となるのです。

ただ今、ご覧頂いたのが、後ほどデモンストレーションで見いただく MoSaIC II のシステムでございます。我々は約2年間、デジタルミュージアムというものを作り上げようということで、DMCの中でディスカッションを重ね、どういうものが新しいデジタルミュージアムなのだろうかということを考えてまいりました。

今日の講演の流れですが、まず最初に DMC におけるデジタルミュージアムとはどういうことを考えているのかというところを軽くご紹介して、アナログ的デジタル技術とデジタル的デジタル技術についてお話できればと思っております。というのは今までのお話にあったこれまでのアーカイヴィングというものは私のなかではアナログ的デジタルアーカイヴィングなのではないかと思っております、デジタルの本当の良さ

というものを活かしきれてないのではないかというふうに考えたからです。実のところ、我々は、この研究活動を通して、デジタルで何ができるのだろうか、デジタルだからこそできることは何なのか、それを深く探求することによってデジタルミュージアムのあり方を考えていこうと研究を進めてまいりました。

さて、DMC におけるデジタルミュージアムです。このシンポジウムの最初に松田先生のほうからご説明がありましたけれども、戦略的研究基盤形成事業の補助金をいただきまして、デジタルミュージアムを作っていこうということを目指しております。そこに書いた目的は、文化財が持つ多面的かつ多様性のある文脈をグローバルに共有して、文化財コンテンツとして統合的に表現していきましよう。モノひとつ、何か文化財があったときに、それはいろんな側面を持っています。そのそれぞれの側面が次の何かの新しい展開を生むのではないか、というのはさきほど上崎さんからのご説明にあったところからもご理解いただけるのではないかと思います。そういった新しいデジタルで何ができるのだろうかといったときに、それをコンテンツとシステムの両側面から研究活動を進めていきましょう。ただ単に情報工学科の我々がこんなの使ってくださいよ、と言うのではなくて、アプリケーション側、コンテンツ側、実際使う側が何を要求していて、それを満たすために技術側から何ができるのか。どこがボトルネックになっていて、問題が解決しないのか。そのあたりを上と下の両方からサンドイッチで攻めていくというのが我々の研究のやり方です。今までのアナログで作ってきたミュージアムのやり方を否定するわけではありません。我々が目指しているのは、アナログはアナログの良さがある。アナログの良さをデジタルで再現するのではなくて、デジタルはデジタルの良さを追求しよう。その2つを組み合わせることで補完し合う表象ができるのではないか。これが我々の考えているデジタルミュージアムのゴールです。

冒頭でアナログ的デジタル技術という言葉を出しましたが、では、デジタル技術って何でしょうか？ ちょっと技術的なことをかじられた方だとすぐこういう話が出てくると思います。CDだと44.1キロヘルツでサンプリングされていて、いい音が聞けます。プ

口になってくると 48 キロヘルツを使います。映画だと 48 キロとか 96 キロヘルツを使います。更に 192 キロヘルツもあります。チャンネル数も 2 チャンネル 5.1 チャンネル、22.2 チャンネル。そういうふうなところがデジタルで表現するということなのかなというふうに思われている方もいらっしゃると思います。映像はどうか。ここには 4K のプロジェクトが入っておりますが、空間解像度 SD、HD、4K、8K。時間解像度が 1 秒間に 24 フレーム、1 秒間に 30 フレーム、1 秒間に 60 フレーム、120 フレーム。このようにどんどん、どんどんビット数を増やしていくことによって達成するのはいったい何だろうかと考えたときに、それはあくまでアナログ情報のデジタル置換によって、それで精度を上げていこうということにしかかなり得ない。要するにゴールは究極のアナログですよ、ということにしか、このベクトルの方向を進めていっても出てこないのではないかと考えるわけです。これが私の今日説明するアナログ的デジタル技術という言葉になります。使っている技術はデジタルです。でもゴールになるのは究極のアナログを目指します。これがアナログ的デジタル技術です。

逆にデジタル的デジタル技術というのは何でしょうか。それはアナログにはできないことを実現するためのデジタル技術と考えております。すなわち新しい価値観がそこになれば、それはデジタル的デジタル技術ではないということになります。非常にシンプルな例です。最近だと船橋の劇場に Dolby の Atmos という装置が入りました。これはすごく高い臨場感の音響を実現できる装置です。何をやっているかということ、こんなことができます。A さんのお家で録音したバイオリン演奏をあたかもサントリーホールで演奏しているかのように B さんの家で再生する。これはアナログではできません。デジタルだからこそできるようなテクノロジーです。A さんがお家、何かちょっと特殊な三角形のお家に住んでいたとしましょう。そこでバイオリンを演奏します。音響特性はこの三角形の家によって決まります。そこで普通録音したものを再生すると、この三角形の家の音響特性に再生する B さんのお家の音響特性が加味されてよくわからない音になって再生される。いい演奏だったら悪い演奏になりますし、悪い演奏だったらいい演奏になるかもしれませんが、そういうふうなことになります。しかしながら、ここにワンクッション入れてやることで、新しいことができますよというのがこのデジタル的デジタル技術のおもしろいところです。A さんのお家で録音

したときに合わせて A さんの家の環境、音響特性を保存しておきましょう。それを使ってある種正規化したこの A さんのバイオリンの演奏を持っておきましょう。そこにサントリーホールの音響特性を重ね合わせてやって、更に再生する B さんの家の音響特性を重ね合わせてやることで A さんの家で録音したバイオリン演奏がサントリーホールで演奏しているように B さんの家で再生されるということが出来る。これがデジタル的デジタル技術です。

まったく新しいことを言っているわけではありません。実際にさきほど申しました Dolby の Atmos もそうですし、他の会社からも出ております。これは、デジタル信号処理という分野です。身近なところだと、みなさま方がお使いのスマートフォンは電波で通信しているわけですが、それにも基本的にはこの考え方が使われています。このデジタル的デジタル技術を実現するときに何が必要になってくるかということ、この真ん中の正規化した表現の構成になります。この正規化のためには A さんの家で取ったこの情報を捨象しなくちゃいけない。ある部分を切り捨ててやることによってこれが作り出されます。そして削り落とされたエッセンスだけを集めて合成することによって新しい可能性がここで生まれてくるというのがこのデジタル的デジタル技術になります。この捨象と合成のバランスが取れていれば新しい価値や利便性が生まれてくるわけですが、取れていないとうまくいかないということになります。うまくいっている例がさきほど申しましたように音響設備であったり、通信、携帯電話や無線 LAN 等で使われている技術です。

ここまでは大まかなデジタル技術のご説明だったわけですが、デジタルアーカイヴィングの話に戻りたいと思います。現在行われているデジタルアーカイヴィングの考え方というのはデジタルデータを将来に渡って保存して活用していきたいということになります。そのときに必ず出てくる話にこのような話があります。デジタルはちゃんと保存さえしていれば経年劣化しないのでスーパー安全保管庫です、という話です。ストレージメディアの寿命が長くなれば、スーパー安全保管庫ができますよ、安くなりますよ、と。保存データからもれなく該当データを見つけたい。さきほどすごく俗人的になっているというお話がありましたけれども、図書館の場合だとスーパーライブラリアンなり、映像制作の現場ですとテープ等を管理しているスーパー人間、間違いなくスーパーライブラリアンを目指している。これが現在のデジタルアーカイヴィン

グの考え方なのではないかと考えるわけです。でも、ここで冷静になって考えていただきたいのは、これは何かデジタルになって新しい価値を生み出しているのでしょうか。ただ単にアナログのアーカイヴィングの考え方をより一歩進めた、もしくはより二歩進めた、より十歩進めただけにしか過ぎないのではないかと。それがこのスーパーの意味です。スーパー安全保管庫。スーパーライブラリアン、これを実現しましょう。これはアナログ的デジタル技術の考え方なのではないでしょうか。

デジタルでデジタル的にデジタルアーカイヴィングをするということは何なのかということを考えない限り、デジタルでのアーカイヴィングの利活用というのは進まないのではないかと。それが我々の根源的な問いなのかなと考えています。そうすると、アーカイヴィングということは、何か情報を使うのですよね。使うってどういうことですか、ということをもう一度見直してみようといったことで情報処理のプロセスというものを基本に立ち返って考えてみようとしたわけです。

さきほど御三方、みなさまいろんな表現されましたけれども、アーカイヴズというものは基本的に保存された情報を利用するためにあるのです、ということをおっしゃっていました。人間の情報処理のプロセスというのはどういうところにあるかということ、新しいモノを見つけます。それは新しい物かもしれませんし、新しい事かもしれません。でも新しいモノを見つけます。そして、その見つけたものが今まで自分が知っていたモノと比較して何が一緒で何が違うのかという特徴付けを行います。その特徴付けに基づいて分類という作業を行って、頭の中を整理するわけです。毎日、毎時間、毎分、毎秒、我々はこれをただ単に繰り返しているだけです。すなわちアーカイヴズというのは情報を利活用するものだと考えるのであればプロセス、この発見、比較、分類のプロセスの保管庫がアーカイヴズそのものなのではないかと思うわけです。

この発見、比較、分類は誰が行うのでしょうか。それは我々一人ひとりが実は行っているということになります。個人が発見、比較、分類を繰り返している。そして、人類が持っている知恵なり学問といわれているものは何かということ、この各個人が繰り返し行っている発見、比較、分類の繰り返しを共有することによって、あ、この考え方は誰々さんと一緒だよ。何パーセントぐらいの人と一緒にだよ。大多数の人と考え方が似ているよね。そうしたらそれはコンセンサスとれ

るよね。そういうことによって知の体系から構造化、構造化された、もしくは体系化された知の上で新しい知の創造ができるのではないのでしょうか。さきほどメタデータの鮮度という話がありましたけれども、分類が知の体系構造である限り日々変化しています。それは学問上教科書で変わってなくても我々の個個人の生活の中で日々変化しているのではないのでしょうか。この細かい動きをきちんととらえて吸い上げていくことができれば、今までにない新しい情報処理のプロセスができるのではないかと考えたわけです。

では、現在の IT 技術は何かできるのでしょうか。今の IT 技術ではできないのですかといったお話になると思います。現在のウェブサービスは、ざっくり言うと、あるデジタルデータがどこかにあったときに、それを取ってくることしかできないのです。ウェブブラウジングはこれを繰り返しているだけです。あそこにある何々が欲しいといったら取ってくる。表示する。それを繰り返しているだけです。何も新しいことはありません。我々が知っているのは、情報というものが世の中にいっぱいあるということだけを知っていて、その中の何かを取ってきて、見て、喜んでいる。それだけです。さきほど申し上げました情報処理のプロセス、何も実現していません。発見もしてくれません。比較もしてくれません。分類もしてくれません。それが現在の IT の技術です。

検索技術は何でしょう。検索技術は何か見つけてくれているような気がするけれども、Google、Yahoo、Bing、いずれも発見技術とは違うというのが私の認識です。キーワードという言葉、普通に使われていると思いますが、冷静になって考えていただきたいのです。キー・ワードです。鍵なのです。鍵を持っている人だけが入れるのです。逆に言えば鍵を持ってないと、その情報には絶対アクセスできないということです。Google が出しているのは…すいません、Google が悪いわけじゃないです。Google、もしくは Yahoo は彼らの独自のアルゴリズムに基づいて、独自のルールに基づいて、独自の技術制約に基づいて、独自のビジネスモデルに基づいてある一面だけを切り取って提示しているだけにすぎません。すごく強引な言い方をすると、機械的で一義的な発見を一方向的に押し付けています。誰も答えを知らないからです。みなさんがグーグルした結果を、みなさんは正しいと思っているかもしれないけれど、本当の答え、もっといい答えはインターネットのどこかにあるかもしれないのです。日々、検索をして出てくる結果はグーグルによって書き換えら

れています。正確に言うと削除されたりしています。リクエストに応じて。それは一方的な一面的なものです。言ってみればドラえもんの世界でいうとジャイアンみたいなものです。俺は強いのだ。だからお前はこれだ、と言われて、それをみなさんが納得しているだけです。すごく精度がいいと思われているかもしれませんが。しかしながら、たとえば Google 翻訳、Google の例ばかり出して申し訳ないですけど、〇〇翻訳ってありますよね、ウェブで。日本語と英語をきれいに翻訳できるのでしょうか。それに対してみなさんはノーとおっしゃいます。基本的に同じ技術しか使われていないのです。なぜグーグルの検索を素晴らしいものだと言えるのでしょうか。もっともっといいものがあるのではないのでしょうか。

最近車を買いたいと思ひまして、M社のAという車なのです。4月ぐらいに「出る」という噂が出てきて、それからインターネットで調べる日々が始まりました。最初は「出る」という噂しか出てこないわけですね。どんなモデルが出てくるかもわからない。それを日々調べるわけです。適切なキーワードもわかりません。誰がその情報を持っているかも知りません。何が得られたか。そこでわかった知見です。Googleは毎日結果を更新してくれない。たとえばある日「あ、新しい記事見つけた！」と言って見つけた記事を見ると、その記事が書かれた記事は結構前だったりするわけです。最新のものではありません。出てきた結果に普遍性や信憑性があるわけでもありません。他にいい情報があるということについても、他にいい情報がないということについても何も保障してくれません。ジャイアンが一方的に言ってきた内容を受け入れているような状況において、現在の検索技術を用いて学問をすることはできるのでしょうか。ある1つのモノを見るときに、一面だけしかグーグルは提示してくれないとするならば、それは学問にはなり得ないのではないのでしょうか。そこのある、自分が知りたいことに対して多面的な理解を深めていくことが学問であるとするならば、それを今は実現できていないということになるのではないかと思うわけです。

現在のデジタルアーカイヴズで使われている技術は基本的にウェブで使われている、みなさんが毎日お使いになっている技術と大して違いはありません。したがって、今のデジタルアーカイヴズの技術で学問するのは非常に危ういのかもしれないということになるわけです。キーワードを知らなければならぬわけです。そこに大きな矛盾があります。自分が知りたいこ

とを自分は知らないにも関わらず、自分が知るためには知っていることを求められているわけです。そして検索エンジンの結果は多かれ少なかれ一面的です。物や事の本質的なところに迫る、それには多面的な理解が必要です。しかし、その多面的な理解には現在のITテクノロジーは通用しない。残念ながらそれが現状です。したがって、新しい発見につながるデータにも辿りつかないし、結果として保存された情報の利活用が進まないということになります。ここで私は自分で資料を作っていて気づいたわけです。今のデジタルアーカイヴズの技術がアナログのデジタル技術であるとするならば、きっとアナログでもこれがあまりできてないのではないかと思いました。そう思いながら資料を作って、さきほどお話を聞いていると、アナログのところでもなかなか利活用が進んでいないということに、「なるほど」と思ったわけです。とするならば、アナログで今アーカイヴズという言葉があって使われてはいるけれども、本当であればそれは何も意味をなしていなかったのかもしれない。それを本質に捉えて何か実現できないだろうか。それをデジタルで挑戦してみるのはいいのではないかと思ったわけです。

DMCとして何を試みているかということです。物、事の本質に迫るためのデジタル的デジタル技術ということになります。情報をさきほどのデジタル信号処理と同じように考えただけならば幸いです。Aさんが持っている情報空間もしくはコンテキストと言ってもいいかもしれません。そこから捨象することによって正規化されたAさんの情報分類が得られる。複数の人の正規化された情報を組み合わせることによって新しいサービスが得られる。それが今我々の考えているシステムの基本的なコンセプトです。この正規化された空間、これを我々はカタログと呼んでおります。捨象するということは何かを捨て去ることです。逆に言えば何かだけを残すことです。カタログで実現しているのは情報分類、何かの情報があって、それをグループ化するなり分類するなり、それを構成する関係を、関係があるかないかだけの1、0に絞ってみました。ギリギリまで削りました。合成というのは何かというと、この削って、削って、削って、残ったこのモノとこのモノには、この情報とこの情報には関係があるということ積み上げていくことで何か出てくるのではないかというアプローチです。

ちょっと具体的なイメージをお見せしようと思ひます。これがカタログによる正規化です。たとえばここに3人、なんでアインシュタインは舌を出してい

るのかいつも謎なんです、3人がいます。物理学者の集合です。こちらにも3人います。こちらは愛猫家の集合です。捨象するとどうなるか。この3人が1つのグループだよ、この3人が1つのグループだよということだけが残ります。それが愛猫家であるか物理学者であるか、そんなことは捨て去ります、まずは。その正規化されたこの関係だけを使って合成してみましょう。そんな情報を捨てたら分からなくなるに決まっているじゃないかと思われるかもしれません。でも、どうでしょうか。これが今我々の作っている、最初にビデオで観ていただいた MoSaIC II というものです。関係を可視化したところ、一本、一本は捨て去られた、ただ単に関係があるかないかだけの線に過ぎないわけです。でも、それを重ね合わせていくと、「あ、ここになにか関係があるな」というのが見えてくるわけです。これは情報というモノを俯瞰的に見て、マクロな視点からデータが持つ、情報が持つ関係を列挙して提示する。そうすると多面的な視界が開けるのではないかというふうに考えたわけです。

この MoSaIC、なぜ MoSaIC というかということ、一番上にありますように Museum of Shared and Interactive Cataloguing というのをつなげて MoSaIC と呼んでおりますが、もとはブラウザの MOSAIC でございます。今ここでは、一番単純化した場合、合成するとこういうふうなので新しい情報が見えてきますよといったご説明をしましたがけれども、我々としてはどこまで情報を捨て去ればいいのか、ある程度残したほうが便利なんじゃないのかなという、一抹の不安感もやはりございまして、こんなトライアルもやってみました。合成に付属情報を用いるということです。たとえばこれ、世界中の都市をエリアごとに分けて矢印で書いたらこんな感じになります。これがそれぞれのノード、点になっているところが位置情報を持っていたらどうなるか。こんな感じになります。これはこれでおもしろいかなということを思っております、うまくこの捨象と合成を実現していくことで新しい知の世界がデジタルで到達するのではないかと考えている次第です。

そういった世界をデジタル情報処理と呼んでみました。情報処理というと、普通、高校の授業とか、年齢にもよると思うのですが、そんなのデジタルに決まっているよと思われるかもしれません。しかし、情報処理とは、本当はアナログでもデジタルでも両方発生することだと私は思っています。そして、そのデジタルで情報処理をするということはどういうことなのかと

いうと、実はさきほど申しましたようにデータ間の関係の有無だけを保存して正規化してやる。正規化してやると有無だけです、0か1です。コンピュータが大好きな0か1。そうすると、あ、この2つはたくさん関係があるねとか、この2つはあんまり関係がないね、というのが数としてコンピュータは数え上げることができるわけです。文字通り。そうすると、このコンピュータの得意なところを活かしてデジタル的デジタル技術、多面性の提示やもしくは物事の本質を知る、正確な情報を知るといったことができるのではないかと思うわけです。

情報の扱い方。よく言う言葉に「知る人ぞ知る」という言葉があります。でも、こういった世界が来ると、誰もが知り得る情報です。今世界中のウェブサイトにはどんな情報があって、どういうふうなリンクが貼られているかは Google や大手の検索エンジンしか知りません。そしてもう我々が手の届くところにはないわけです。なぜかというとそのだけのデータを集めるためには膨大なリソースが必要だからです。ストレージリソース、コンピューテーションリソース。それを持ち得なければこの情報を我々が知ることができないのが現状です。でも本当は、情報というのは我々一人ひとりが判断して答えを出していくものではないでしょうか。個人が情報の価値の判断をする。そういったデジタル情報処理の時代に突入できれば私は嬉しいなと思っております。

本日、このあと、上でデモンストレーションをご用意しております。実際に関係だけになった世界でどんなことが見えてくるのか。実際に触れていただくとうご理解いただけるのではないかと思います。我々も最初は、最初の言い出しっぺの私も不安感があったわけです。こんなに捨て去っていいのだろうか。頑張ってメタデータを付けて、それを大活用すればいいのではないかと思いましたが、でも、DMC のみなさんが「やってみよう」ということでやってみた結果、非常におもしろい世界観ができてきております。是非とも自分の手で見て触れて感じていただければと思っております。以上です。