

知のデジタルアーカイブに関する研究会（第5回）議事録

1. 日時：平成23年8月24日（水）10：30～12：30

2. 場所：総務省8階第1特別会議室

3. 出席者（敬称略）

（1）構成員

新麗、安達文夫、入江伸、植村八潮、大内英範、大場利康、岡本明、小川恵司、加茂竜一、杉本重雄、武田英明、田中久徳、田良島哲、常世田良、栗原浩（鳥越直寿構成員代理）、宮澤彰、盛田宏久、山崎博樹、風間吉之（八日市谷哲生構成員代理）

（2）総務省・文部科学省・経済産業省

（総務省）佐藤政策統括官、阪本大臣官房審議官、黒瀬情報流通行政局情報流通振興課長、松田情報流通行政局情報流通振興課統括補佐

（文部科学省）高尾文化庁文化財部伝統文化課文化財保護調整室専門職

（経済産業省）望月商務情報政策局文化情報関連産業課課長補佐

4. 議事

（1）Linked Dataによる博物館情報の共有・再利用

（2）クラウドの視点から見たデジタルアーカイブ

（3）文化財の情報化とバーチャルリアリティ（VR）公開

（4）デジタルアーカイブの構築・連携に係る調査研究について

（5）その他

5. 議事録

【杉本座長】 おはようございます。それでは、定刻となりましたので、「知のデジタルアーカイブに関する研究会」第5回会合を開催いたします。

本日はご多忙の中、ご出席いただきまして、まことにありがとうございます。

本日は、神門委員、丸山委員、水谷委員が所用によりご欠席と伺っております。

まず、議事に入る前に、配付資料の確認を事務局よりお願いいたします。

【松田情報流通振興課統括補佐】 お手元の資料でございますが、議事次第がございまして、その下に座席表、資料知5-1が、武田構成員の「Linked Dataによる博物館情報の共有・再利用」、資料知5-2が「クラウドの視点から見たデジタルアーカイブ」、新構成員の資料でございます。資料知5-3が、加茂構成員の資料でございまして、「文化財の情

報化とバーチャルリアリティ公開」、資料知5-4、「デジタルアーカイブの構築・連携に係る調査研究について」ということで、5-4は、5-4-1、5-4-2、席上配布のみでございますが5-4-3として、「デジタルアーカイブ構築・連携のためのガイドライン（たたき台）」。最後に参考として、第2回の「検討アジェンダ（案）」、こちらをお手元にご用意しております。

【杉本座長】 どうもありがとうございました。

それでは、議事に進みたいと存じます。

本日は資料の順番で、武田構成員、新構成員、加茂構成員ということなんですけれども、新構成員に先をお願いして、それで武田先生はその次、2番目に入っていただくということで、ちょっと順番をスイッチさせていただきたいと思います。

じゃ、新さん、よろしくお願いします。

【新委員】 I I Jイノベーションインスティテュート、新でございます。よろしくお願いいたします。

今日は、「クラウドの視点から見たデジタルアーカイブ」ということでお題をいただいているんですけれども、前回、災害とデジタルアーカイブでも少しクラウドについてお話しさせていただきましたので、定義とか最初の話は今回はなしで、もう少し技術的な視点とか、もしかしたら皆様がクラウドに関して持っているイメージが技術的にこうですということが少し言えたらなと思ひまして、準備させていただきました。

まず、そのクラウドとはというところは前回少しお話をさせていただきましたので、クラウドの種類ということで今日は始めさせていただきたいと思うんですけれども、一般的にクラウドといいますが、ほんとうにさまざま定義もありますし、種類もあるんですが、ごく一般的な定義として、Software as a Service、S a a Sと言われるものと、それからPlatform as a Service、P a a Sと言われるものと、Infrastructure as a Service、I a a Sという呼び方をしますけれども、そう言われています。実はこれ以外にもH a a S、Hardware as a Serviceであったり、いろんな言い方はあるんですけれども、ごくざっくり言うと3種類と言われています。

それぞれの内容についてですけれども、いわゆるS a a Sと言われるものは、ソフトウェアの部分をサービスとして提供するもので、一般に最も皆さんが目に見えるパターンとしては、S a a Sが多いんじゃないかと思うんですけれども、グーグルのメールですとか、セールスフォースですとか、いわゆるクラウドというものが華々しく出たときに一番話題

を集めたのは、SaaSの部分であると思います。

それからPlatform as a Service、PaaSですけれども、これはもう少し下のレベルになりまして、アプリケーションソフトウェアが稼働するためのハードウェアの部分、プラスOSの部分サービスを、外からインターネットを通じてお使いいただけるような提供にしたものがPaaSとされています。ここにどこまで含めるかはいろいろと議論があるんですけれども、例えばデータベース部分をPaaSとして提供したりという形でも提供されています。

それからIaaSと言われるのが、全くハードウェアの部分なんですけれども、CPUのサイズこのくらい、ストレージこのくらい、インターネットの待機このくらいという形で、本来でしたら組織の中の建物に構築される部分を、クラウドという形でアウトソースでしたり、ホスティングと言われていたものとあまり変わらないんですけれども、インターネット経由で提供できるようにした形をIaaSと呼んでいます。

最初のSaaSのほうはGoogleとかと申し上げましたけれども、Platform as a Service、PaaSとされているもので、多分皆様がよくお目にされるとしたら、Microsoft as a Service、それからAmazon EC2はIaaSに当たるとされています。それから一般に日本のクラウド事業者が提供されているのは、IaaSと言われる部分が多いと言われていて、IaaSを提供しながら上に今までどおりの独自のサービスを構築して、提供して、クラウドとしてサービスしているという形が一般的には多いと分析されています。

それからクラウドサービスの形態ですけれども、これはちょうどガイドラインの案のほうにも出ていたんですが、プライベートクラウド（オンプレミス）とって、これは従来どおりで、組織の中の建物にクラウドのような形のシステムを構築して、それを使うという形です。これは今までの組織内のシステム構築と大きな差はないように見えるんですが、わりと大規模な組織で、各部署ですとか部門ごとにそれぞれ資産を持っていたものを、1カ所に集めてクラウドとして使うというような形態が多いので、一般的には大規模な組織の方が構築されているパターンが多いです。これは今までと同じようにやりますので、詳細なカスタマイズが可能です。

ハードウェアの保有は全く提供形態によると書いているんですけれども、組織内に置かれる場合もありますし、これ自体をデータセンターに置かれるパターンもありますが、一般にハードウェアは組織の方が保持したままという形になります。

それと大きな対極としてパブリッククラウドというのがありますけれども、これはシステム全体をクラウド事業者のほうを利用するという形です。パブリッククラウドというのが、今一般に言われているほうのクラウドだと思いますけれども、ある程度定型化されたものを選んで、組み合わせてご利用いただくという形です。

最近特に日本でわりと要求が高いものとして、ハイブリッドというのがあります、まさにプライベートとパブリックを組み合わせる形なんですけれども、それはまたシームレスにつながる形で、どうしても外に出せないデータについては組織内に置き、それもクラウドとして利用し、ウェブデータや公開するデータのように、パブリックでもよいものはパブリックを利用しながら、それを組み合わせるという形でも出てきています。オンプレミス、プライベートになりますと、やっぱり非常に高価格になりますので、それとパブリックとを組み合わせ、要件に応じて使い分けるといった形が結構出てきています。

デジタルアーカイブにおけるクラウドということで、ここで会社ではなくてこれは私の私見なんですけれども、申し上げたいと思いますが、今々のクラウドの技術を考えますと、まだサービスとしてはあまりデジタルアーカイブに特化した形では出てきていないだろうと思われまます。もちろん皆さん調査なさってみて、どうも使えないと思われることが多いと思うんですけれども、そういう意味では、デジタルアーカイブという視点でクラウドのサービスはつくられておりませんので、今良いサービスというのは、おそらく選んでいくと難しいのではないかと思います。

では、何のときにクラウドが使えるかということですが、これは最初に分類させていただきましたSaaS、PaaS、IaaSの中ではもうほとんど、IaaSかPaaSのところ、インフラとして今コンピュータを買って構築されたシステム、その部分だけをクラウドとして外に出して使うということで、ストレージサービスとバックアップサービスというのが一番今使えるものではないかと思います。

クラウドとしてのメリットとして必ず言われますのが、初期投資がなくなるということと、それから4年、5年ごとシステム入れかえという形で、これまで5年後に使えるだけのストレージを先に確保しておくことをされていたと思うんですけれども、その分の余剰資産はなしで、増えた分だけその年、その年で買い足していくことができることです。それから運用コストの部分で、ハードウェアの運用というのはクラウド事業者がやりますので、その分のコストとか、あるいは24時間体制でのバックアップサービスとか、サポートサービスということは、違う形でつけるので、一般的にはコストとしては下がる

というのが現状です。

ストレージサービスについてですけれども、これもオンラインストレージとか、諸々いろいろあるんですけれども、ざっくり言うとデータをクラウド上に置くサービスにはなりません。特徴としては、必要に応じて容量を増やしたり減らしたりできるということと、それからディスク単価がどんどん下がっていますので、大体例えば今年必要な分だけ買って、来年は容量が増えたんですけれども、その増えた分を同じ価格でサービスしますという形で、だんだん値下がりしているというのが現状ですので、そういう意味では初期投資とプラス維持費というのも、今下がる傾向にあることは確かです。

ここから先、私は何回もクラウド事業者の運用方針というのを書かせていただいているんですけれども、やり方がさまざまですので、いろんな形でよく検討される必要はあると思います。まずストレージサービスについては、個人利用としてよく見える、写真の共有サービスですとかメールサービスということは、使い方に特化した傾向がありますので、いわゆるSaaSの部分で、このままストレージに使うのはちょっと難しいだろうと思われれます。

それ以外にストレージエリア、これはほんとうにビット配列を置くというだけのもので、例えばメタデータですとか、マイグレーションとかは全く考えられていないサービスですけれども、そのエリアとして提供してカスタマイズするというサービスは、結構出てきています。それと別途ストレージはそれなりに需要がありますので、共通APIというのをつくっていく動きもあります。

それからバックアップですけれども、これもちょうどメーリングリストで大変皆さんも話題だったので、どういう情報をご提供したらいいか、少し悩んでいるところがあるんですけれども、私として今数字を持っているわけではないんですが、テープですとかディスクとかにバックアップをとっていくというシステムが、運用上それほどコストが下がるわけでも信頼性が上がるわけでもないというのが事業者側の印象としてあります。データコピーに置いてあったほうが、結局復旧がしやすいので、テープやディスクが思ったほどの信頼性を保持してくれないので、最近ディスクの低価格化もあわせて、データコピーにしてテープやディスクをやめて、ストレージとしてコピーを持つというような使い方が増えてきています。

こうしますと、媒体の保管ということがなくなりまして、クラウド上で地域分散ということをする形でバックアップがとれるようになります。これもストレージと同じですけれ

ども、容量に合わせてリソースを増減しますので、ほんとうにこれはコストとのご相談ですけれども、テープメディアあるいは再生機器の保管をするよりは、いろんな意味でコストとしても見合ってくるのではないかなと思いますが、私として今詳細な検討結果があるわけではないので、傾向として申し上げます。

それから、頻度とか形式とか復旧までの時間要求というのはもうさまざまで、それぞれのデータの性質によって、いろいろと考え直さないといけないことは確かではあります。

それから、クラウドというと必ず言われますセキュリティなんですけれども、これは語り始めると、1時間、2時間で終わらないほどいろいろありますけれども、ごく一般的なお話としてさせていただきたいと思います。クラウドはやっぱりセキュリティが危ないとよく言われるんですけれども、クラウドをどうとらえるかではあるんですが、一般的な言い方として、クラウドが公開とイコールだと思っていると、必ずしもサービスはそうでもないということは申し上げておきたいと思います。

プライベートクラウドという形で申し上げましたけれども、プライベートクラウドといわれるような、クラウドを組織の中、建物の中に置く場合もありますし、事業者のデータセンターに置く場合もありますけれども、そこと各組織の間のネットワークもプライベート化して、安全にアクセスできるような形を提供するというのは、わりとよくつくられておりますし、もちろん暗号化でデータを守るということも、機能として幾つも見られます。コンテンツに入れることもできますし、ネットワークに流れるデータ自体を暗号化するという技術も今進んでいますので、そのあたりをきちんと守れるようにはなっています。

それからアクセス管理ですとかID管理ということも、お客様ごとに違うものを必ず準備してアクセスできるようにしていますので、全く公開ということもないですし、同じ場所に置かれるということも普通はないようにつくことは、もちろん可能とお考えいただければと思います。

クラウドもセキュリティによりますけれども、クラウド特有の問題というのは、ごく一般的に言われているものから言うとそれほど多くなくて、ほんとうはITセキュリティ、アクセス管理だったり暗号化であったりということの延長で、大体同じように、この延長上としてクラウドへの移行というのは考えていただけるのではないかと思います。

ただしなんですけれども、セキュリティに関してどう公開して、何を対策しているかというのは、まだクラウドに限らず、ITシステム一般として確固としたガイドラインというのは、全くないわけではないんですけれども、十分にはないと言えらると思いますので、

ニーズに合わせて選択する必要というのはもちろんあります。

それからセキュリティというか、信頼性に関してなんですけれども、物理的なデータの置き場所というのは、一般にデータセンターのような大規模なシステムをつくりますと、電源管理から地盤から耐震性からということで、いろんなことを考えますので、どこが安全というのは言いにくいですが、一般的にはかなり安全なところにつくってはおりますので、ディスクがどこからか落ちてしまうとかいうことに関しては、少しリスクは少な目であろうとは思いますが。

それから入館に関しても非常に厳しいですので、物理的なセキュリティ、ほんとうに泥棒がやってくるようなことに対しても、かなり強くできていることは確かだと思います。

それからセキュリティとか信頼性への対処として、どんなことが考えられるかなんですけれども、一般にデータがなくなってしまう、消えてしまうということは、先ほども申しましたようにバックアップのいろいろな体制はとれますし、それからアクセス回線の盗聴ということに関しては、もちろんプライベートネットワークを最近では組むことで、プライベートネットもいろんな形がありまして、インターネット上のプライベートの場合もありますし、回線自体をもっと安全に守るという技術もありますので、それは要求に応じていろんな形は可能です。それからもちろん暗号化というのもあります。

それから1カ所に置いたものが盗難されるとか危ないということで、最近秘密分散、あるいは割符という形で結構ありますけれども、それもシステムの要求に応じてつくことは可能ですので、クラウドのセキュリティ的なことを高めていくことは技術的にはいろいろありますので、どういうものが必要かということを見直していただくと、それなりの対処がいろいろ考えられるだろうと思います。

これは事業者側として申し上げるわけではないんですけれども、ただ、技術的な実現方法ですとかサービスとして出ているものはいろいろな形がありますので、要求事項をきちんと見直されることと、それに応じた事業者を選択していただくことは非常に大事だと思います。

ということで、ちょっとまとめにはならないんですけれども、デジタルアーカイブから見た現状のクラウドとしては、やはりサービスレベルとしては不足であろうとは思いますが、それはこれから十分時間をかけ、要求を整理して、システムとして発展する必要がありますので、デジタルアーカイブというサービスを考えた上で、どれだけのシステムが必要かということは、ぜひこの場で検討いただけるとよろしいのではないかと思います。

それから、いろんな意味でシステムというのが変わっていく場合に対応する形として、最近ではアプリケーションインターフェースの統一と言われていまして、これが標準化、定型化されていくことは、ある意味システムの寿命を延ばすというか、システムが長く同じように使えるという形になりますので、アプリケーションインターフェースの定義というのは非常に大事だろうと思われまます。

それから、クラウド化によってシステム設計をどう変えていくかということなんですけれども、まずリソースとして余剰リソースを考えるとすることは、これからはクラウドの場合は考える必要がなくなりますので、リソースの考え方を少し変えて設計をしていただく必要があると思います。

それからシステム管理が随分変わると思うんですけれども、ハードウェアの故障というのは一般的には事業者がやりますので、普通サポートサービスとかの形は不要となります。それからサービス部分についても、クラウドの事業者から買う部分に関しては管理は不要になってきます。この場合に先ほどのセキュリティにもなりますけれども、例えばOSでのセキュリティホールが見つかりましたというような対応に関しては、これも事業者側がやりますので、セキュリティレベルを最新に保つということは、クラウド事業者側が責任を持つということにはなってきます。

それから人材に関しても、この会議で随分検討として上がっていたと思うんですけれども、クラウドになると実は必要とされる人材が大きく変わるとというのが、一般的に言われるとはちょっと言いがたいんですけれども、あまり知られていないんですが、わりと重要な話です。

これまでシステムのハードウェアであったり、ソフトウェアであったり、システムの組み方であったり、設計に関してと、それからそれを運用し、管理して設定をするというのは、ほんとうにシステムに非常に詳しい人材が必要だったところから、運用とか管理とかというのは、もう事業者任せという形になりますので、人材としては企画をしたりですとか、あるいは要求をまとめたりということに関してできる人材のほうが必要になってきますので、コンピュータエンジニアではなくて、やっぱり企画のできる人材が必要になります。

これは考えてみますと、システムの細かいことは任せることにして、それぞれやはり各組織で必要なものを整理していただくという、本来の業務に戻られるということにはなりますので、ある意味うまくクラウドで人材を補っていただくということは、考えていただ

けるのではないかと思います。

まとめですけれども、ここ数年のコンピュータとネットワークの技術の傾向とサービスの傾向からすると、クラウド化というのは今始まったものではないというのがありますので、今後も継続はするだろうと思います。

それから、コンピュータとネットワークがそれによって使い方というか、利便性が非常に向上していますし、巨大システムが受けるようになったというのがありますので、この流れはおそらく続くと思いますので、事業者の継続問題というのが別にありますけれども、クラウドがなくなるということはおそらくないであろうと思います。

クラウドのほうの課題なんですけれども、今まさに標準化真っ最中という感じで、言葉の統一からサービスのいろいろな形からというのが標準化がされていけませんので、今どこかのシステムを買った場合に、その移行をするというのはちょっと困難であります。同じような形であればいいんですけれども、どこを同じレベルにするかというのはわりと決まっていますので、どのシステムも同じなんですけれども、まだこの辺はそういう整理がされていません。それからいろんな意味で指標も統一されていけませんので、判断に迷うですとか、情報が多過ぎるということは、今実際あると思います。

ほんとうに今、標準化活動はいろいろな団体がやっています、ハードウェアでもやっていますし、ストレージ周りでもやっていますし、ネットワークでもやっていますし、あるいはシステムでもやっていますし、サービスレベルでもセキュリティレベルということで、非常に標準化活動が活発ですので、標準化とガイドラインの見直しがここ数年出てくると考えているところです。

非常に技術的な話ですけれども以上です。

【杉本座長】 どうもありがとうございました。後でまとめて質疑をする時間はとってあるんですけれども、今この場でクイッククエスチョンがもしございましたらば。よろしいですか。じゃ、どうもありがとうございました。

じゃ、続きまして武田先生、お願いいたします。

【武田座長代理】 国立情報学研究所の武田です。

私のほうでは、最初にオントロジーの話とかいう話もいただいたんですけど、少し私が最近やっていることと、今回ここでずっとアーカイブの話をしていて感じる点を、ちょっと自分なりにまとめました。ポイントは、ここでは主にいかにデータを保存するかという話なんですけれども、私の視点は共有して使うとか、その点を含めてアーカイブというの

を考えてほしいなという意味で、今回の少しデモを交えたご説明をしたいと思います。

こんなことは皆さんに言うのもおかしい、常識的な話ですけど、やっぱり情報というのは使って何ぼであると。要するに幾ら情報をため込んでいても、それはだれかに使われなきゃ意味がない。使うということがまた新しい情報を生むんだというのが、健全な動きであらうと考えています。だから情報というのはやっぱり過去の次の蓄積からまた生まれてくるものである。だから集めて使うというのは、むしろ我々の理知的活動の当然の行為である。

それからまさに情報というのは使われてこそ意味がある。使われない情報というのは、それこそほんとうに私たちに何も価値がない。つまり利用するというはやっぱりそれを公開することである。つくったからそれはどこかに置いておくのではなくて、やっぱり公開しないと使ってもらえない。そうしたら公開されたものを共有できる仕組みというのは、まさに我々の社会、文化の基盤ですね。

それはかつてもちろん、こんなのは別にコンピュータ、ネットワークがない時代から、ずっと出版を通じてとか、物を実際に公開してとか、そういう形でやってきたわけですね。それが今はデジタルの時代になったので、ネットワークの上でこういうことをやると。やっていることは同じなんだけど、ネットワークの時代はネットワークを通じてこういうことをやるというのがむしろ普通になったねと、それが現状だと思います。

そういった中で、この研究会で主にずっと議論しているのは、どちらかというは収集する、それからそれを利用する、つくる、公開する、このあたりに焦点を当てて話をしている。もちろんそれはとても大事なことで、まさにここにいるいろいろな機関の方々が、実際にやらなければ動かない話であります。ただ、前のスライドで書いたように、それはやっぱりつくるだけではなくて、使ってもらわないといけないんじゃないかというのが今回の論点であります。

そこで、一たんちょっと話は切れますけど、ネットワークの世界、特にウェブの世界で、どうやって情報を共有するかという問題に関して、これは随分、もう10年以上こういう議論をしてきました。

セマンティックWebという考えがあります。セマンティックWebというのは、いわゆるウェブ、普通に我々が使っているHTMLというのは、どちらかというは人のための情報であるということから問題が提起されています。これはご存じだと思いますけど、HTMLで書くというのは相当楽なんです。うまく、わりと簡単に書けるので、しかも人間

が読んで意味がある。けども、実はそこに書かれている構造というのは、コンピュータにとってはあまり意味が読み取りやすくないんです。

そこでセマンティックWebという考え方においては、人だけじゃなくてコンピュータも利用しやすい構造を与えようということを目指しています。これはWorld Wide Webをつくったティム・バーナーズ＝リーという人がみずから立ち上げたプロジェクトで、彼は実はウェブをつくる時に、今で言うセマンティックWebをつくりたいと思っていたんだけど、つくったときに結果的には今のHTMLをつくってしまった。なので、もう少し人とコンピュータ両方が理解しやすい情報を書きたいというので、セマンティックWebという方法を提案しています。非常に大ざっぱにしてみると、この3点ぐらいになります。

URIというのは、大体皆さんが使っているウェブのアドレス、URLと同じものだと思っていただいて結構です。ちょっといろいろ違うんですけども。とにかくURIによって物を記述しましょう。これは後でちょっと議論しますが、すごく重要なことなんです。それからメタデータによってデータを記述しましょう。これはここの研究会の方はむしろ当たり前かと思えますけど、逆にウェブの世界では当たり前じゃないんです。むしろメタデータで情報を構造化しましょうということでもあります。さらには、メタデータ同士をオントロジーで構造化しましょうと言っています。この3点が3行で言うセマンティックWebであります。

これは先ほど言いましたように、もう10年以上前にこのやり方を提起していますがけれども、リアルな運用の中でそれほどこれが適用されているわけではありません。セマンティックWebはねえるところはすごく大きくて、ウェブの全情報を構造化したい。でも、なかなか今あるウェブを置きかえるような形にはならないというのが現状でした。

ところが2年前ぐらいですけれども、実はウェブの情報全部を構造化するというのはなかなか大変だ、むしろ困っているところがあると。むしろデータの共有にはどうしたらいいか。そこで実はLinked Dataという考えが出てきました。Linked Dataというのは実は何のことはない、このセマンティックWebの3項目のうちの上の2つだけ、URIによる共通基盤とメタデータによるデータの記述、ここだけに特化して、特にしかもドキュメントでなくて、データの構造化に貢献しようというものです。これがこの2年ぐらい、非常に大きな動きとなってきました。

このLinked Dataというのは通称Web of Dataといいまして、今までのドキュメントがWeb of Documentだったのに対して、これはWeb of Data、データのウェブなんだと。何がウエ

ブなのかというと、R D F、Resource Description frameworkという、ある非常に簡便な言語を使って、まずデータをすべて書きます。1つの記述方法で書きます。それからそれは、外部から参照可能なもの、つまりリンクができる、あるいは自分がリンクである、リンカブルなデータということです。だから標準化されたフォーマットであって、相互にリンクできる、こういうものをLinked Dataと呼んでいます。

下に、ティム・バーナーズ＝リーが挙げた4条件というのがありますけれども、とにかく先ほど言ったように、まず、すべての存在物にはU R Iを与えてしまう。それからU R Iというのは、皆さんよく使っているものですけど、H T T Pで書かれるものにして、ということは外部から参照可能な情報にします。しかもその参照した情報を見たときに、実はそれが何であるかということがわかるような記述も加えます。こういったことを条件としています。

何かこう言うとすごく難しそうなのですが、実はすごく簡単なことです。実はR D Fというのは主語、述語、目的語ぐらい、3つ組、トリプルなんて呼んでいますけど、たった3つで書かれるデータコードなんです。だからミニマムな一番小さいデータコードです。これですべてを書いてしまおうというのが、このLinked Dataの考え方です。2年前ぐらいに、これでデータが書けるんじゃないかという話が出たときに、多くのデータを持っているサイトが、じゃ、自分たちのデータもこのR D FのLinked Dataで公開しようということを始めました。

これは今年の6月ぐらいの状態で、どのぐらいのデータサイトがデータを公開しているかということを示した図です。基本的にはここではオープンに公開するということが前提なので、Linking Open Data、あるいはLinked Open Data、L O Dと呼んでいます。これはちょっともじりですけど、これをこの世界の人はL O D Cloudと呼んでいます。何となく見た感じが雲っぽいからということで、別に先ほどの話のクラウドとは全く関係ありません。

ともかく、これは1個1個がデータのサイトです。これを見ていくと、後で公開するL Cとかいうサイトもあるし、A C M、これはアメリカのコンピュータ学会ですけども、そういった学会とかも入っていたりします。この矢印は、それぞれのデータ間で先ほど言ったリンクがあるかどうかを示します。真ん中にあるのはDBpediaといって、これはWikipediaをLinked Data化したものですけども、例えばそこに自分のデータがリンクしているよというところの黒い矢印が、そっちへ向いた矢印に書かれる、そういうことを意味しています。こんなものが今世界では起きているということです。

特にMLA関係で言うと、国立国会図書館はつい最近LOD化して、今のところ典拠と件名標目がLODで公開されている。多分書誌もそのうち公開されるんじゃないかと思っていますし、アメリカ議会図書館は書誌、それから件名標目等をやっぱり公開しています。大英図書館、ブリティッシュライブラリーも同じようなことをやっています。実は図書館に関してはすごく今大きな動きになっています。もう大手の各国の中央図書館は、これで出すのが当たり前のようになっていっています。

美術館・博物館に関しては非常に少ないです。先ほどのLOD Cloudにもほとんど入っていません。これはいろいろ問題があることだと思っています。いろいろ問題というのは、データ自体の問題もあって、なかなか進まないこともあるんだなと思っています。時間もあられなのでちょっと早口ですいません。

もとの話に戻ってくるんですけど、美術館・博物館のデータというか、我々LODAC Project というので、実は日本でこのLODを普及させたいという活動を今やっています。これは別に分野を限らず、とにかくデータをLOD化するんだということをやっています。これは去年の4月から始めて、最初に取り上げたのは美術館・博物館の情報、あとは地名・地理情報もやっています。ローカル、これは観光情報みたいなのもやっていますし、今年からは生物種情報などについてもやっています。こういったいろんな分野を横断的に、とにかくLOD化するということをやっています。

今日取り上げるのは、その中の去年の4月から始めた美術館・博物館の情報なんですけれども、これは特にメタデータの世界だと思ってください。いわゆる物そのもの、画像とかいうものではなくて、メタデータの世界。メタデータが例えば今どうなっているかというと、ウェブ上で手に入る情報なんていうのは極めて少ないです。しかも各国が独自で維持しています。一部公立博物館・美術館は合同でサイトをつくっていますが、ほとんどはばらばらです。あと、透明性がないというのは、それぞれで設計しています。

じゃ、それをとりあえずLODのスタイルでくっつけてみたらどんなことが起こるだろうとつくったのが、そのLODAC-Museum (仮) というものです。これよりもっとちゃんとしたものは何かというと、ここにいる方は多くの方がご存じだと思いますけれども、ヨーロッパではeuropeanaというプロジェクトがありまして、これはEUの二十六、七カ国の美術館・博物館が、メタデータを集めて共有しようというプロジェクトです。ここにあるように言語も大量に並んでいます。

このeuropeanaでは、とにかくメタデータを集めて検索可能にしています。データクリエ

イドというのは各国を代表するようなサイトで、実際には1,500ぐらいの機関が入っている。こういうことは実際にやっています。実はこのセマンティック、先ほどのLOD版も今試作されています。ちょっとまだ完全に公開されるというわけではないですけども。これがむしろモデルと言えればモデルです。その日本版みたいなのが欲しいねということです。

やっていることは単純で、ひたすらウェブ上にあるデータを集めて、それをある構造に押し込んでつなげたということです。実はつなげるところが一番難しいんですけども、どういうことか。一応我々にとってのメタデータも設計しなきゃいけないので、いろいろなメタデータスキーマーから借りてきました。46個ぐらいの要素で書いています。これはまだ今でも拡張しています。

結局いろんなところから来たのをくっつけなきゃいけないということで、実際にそのくっつける分については、今主にその作者を中心にやっています。作者ノードが一致したら、それは同じ作者だと。じゃ、どうやって作者ノードって決めるのというと、うちの場合だと日本美術シソーラスというという、筑波大学の先生がつくられたそちらをネタに、そこにつなげるという形で作っていく仕組みをつくりました。

ちょっとそれは見ていただくほうが早いですけど、非常に単純なもので、例えばこうやって検索、リンクしても同じですけど、やれば、こうやって出まして、下村観山という方の作品で、これは日本美術シソーラスからとってきたもので、これはWikipediaで、実はここについている作品の一覧というのが各美術館であったものです。前にクリックするとその中の、実際これはどこに置かれているものかというところと横浜美術館です。そういったいろんな各地の美術館のものが。

例えばこの方の作品がどこにあるかというところ、我々の収集した館、今15館ぐらいありますけど、15館のどこかにあればここに一覧として出てくるというものです。そういうものでこの一覧が見える部分があるし、いろんなところから出てきますねというところが見えます。現状、これは14館のときのデータですけども、名前でも合わせてくると、例えば今のところ600人ぐらい持っていて、結構これでやると、結局6万件ぐらいのうちの1万5,000件ぐらいは、作者で同定できましたよなんていうことも出てきました。

これはかなり荒っぽい作業をしているので、間違いも入っていることはあると思います。今のところここで、作品数で言うと6万点ぐらい。いろんなデータを集めてくると、10万データぐらい今集めています。これは今も実は続きでやっています。一応100館ぐら

いをカバーしようと今やっています。

我々にとって、別にこのデータベースをつくりたいことがほんとうのことではありません。むしろこれを使ったときに、どんなことができるかということを見たいのがほんとうのところなんです。例えばそれは、今までつながっていないものがつながったことによって、一つの主題に当たる情報が集められたりとか、あるいはある場所に関する情報が集められたりとか、そういった多様なデータの使い方ができるということを考えています。

例えば、わりとささいな例ですけども、これはこういうデータを集めて、ちょっと遊びでつくった、「仏」と題名にされた作品を持つ神社はあるのかと調べてみると、ちゃんとあります。40件ぐらい出てきました。これは文化財のデータからとってきたのも多いですけど、神社なんだけどちゃんと如来様とかいうのもたくさん持っているところもありますね。

こういう検索もできますし、あるいはある技法の作品を集めてきて、その作品の年代の幅を見ようなんていうのも。これはちょっとフォーマットが変ですけど、例えば国立近代美術館で、これはテンペラという技法を使った作品。要するに技法欄にテンペラと書いてあったやつを集めてきて、その年代をまた集めてきてみたいなのをやっています。そうすると、例えば国立近代美術館では1984年から1912年までの作品があるねということがわかります。

これはRDFを検索するSPARQLクエリというある言語がありまして、これで自由に書くことができます。ここが普通のAPIと違うところで、基本的にはリレーショナルデータベースのSQLをイメージしていただければいいんですけど、SQLと同じような仕組みを、実はウェブ上から操作してこういう結果を自分でつくることができます。実はこのサイトで、ここで検索をかけると、先ほどのユーザー向けにちょっとインターフェースをつくっただけですけど、こう自由に検索することができたりします。こういうことができるような仕組み。

中にはそれを使って、ちょっとしたアプリケーションをつくってみたこともあります。これは横浜の、まさに美術情報じゃないほかの情報、地理情報と組み合わせて、地図、博物館の位置とかいうのと組み合わせてこう出ます。さらには博物館のこの地区の一覧が見えます。さらにはそれを見るときに、例えば横浜美術館で行われるイベントと同時に見ることができたりします。このイベントの情報も先ほどのRDFで同じく書かれていて、同じ仕組みの中で、こういう一種のポータルみたいなものをつくることもできます。

実は先ほどの収蔵品を調べるのと同じような仕組みを、ここで埋め込むこともできる。これは全部データそのものは、RDFという手法で書かれたものを人間用に、いわばマッシュアップして書いたものです。

実はさらに重要なことが、今回一番申し上げたいことは、つまりこれは今まで何となくデータというのは、そういう機関が提供するものだと思っけていますけれども、むしろユーザーだって貢献できるわけです。例えばこの大日如来というのを調べたときに、仏像のデータをたくさん博物館は持っていますけど、意外に細かい情報を持っていなかったりするんです。例えばこれは、文化財データベースにある大日如来像ですけど、これはどこにあるかという、金剛宝戒寺というところにあるんですけど、データはあまりないんです。

例えばそれにデータをつけたい。じゃ、それを調べてみる。データをつけようと思う。それを例えばWikiを使って入れるなんていうことも、一応仕組みとしてつくりました。例えばここにWikiを使ってデータを追加しようなんて思ったとします。ちょっとこの辺は今まだ開発途中なので、あまり格好よくないんですけども、データをつくらう。例えばここで、このある情報に関してデータを追加したいというときに、これは私がつくった仏像に対する構造的なメタデータをつけたいと、Wiki上でこういうテンプレートをつくれるので、テンプレートをつくります。

ここに例えばタイトルとかを入れる。ちょっと調べてみました。この先ほどのやつを調べてみると、実はこれは如来で、しかも大日如来でとか、例とか書いて、姿勢はというところで坐像と書いてある。材質はもとのデータには書いていないんですけども、例えばこのだれかが書いたウェブサイトによると、木で作られていると書いてあるので、じゃ、これは木だと、こんなふうに。

印相はちょっとこれは読みづらいですね。こうかなとか、こんなふうにして入れていくことが、事物はあるかなとかつって入れると、これがまた一つのデータとなって、これはWikiのほうに記録する。で、Wikiでつくられたものもやはり先ほどRDFで書かれていて、例えば同じ大日如来を持つというものを、幾つあるかなと調べたりする、こんなこともできるようになる。こういったユーザーがつくったデータと結合することもできる。こういうことができるというのが、やっぱりLODのおもしろいところだと思っています。

まとめに入りますけれども、結局Linked Dataというのは情報共有の新しい方法で、メタデータはその意味ではより重要な役割を期待されています。情報の蓄積・公開というのは、共有・利用と一緒にあって価値があるんじゃないか、ここが私の今日伝えたかったところ

で、最後のWikiでつくった例を見せたように、ああやってデータが公開されていれば、それにさらに例えば仏像に対する詳細な情報をつけてみよう、あるいは各絵画でも、この絵画のモチーフになっている場所はどこだろうかとか、そういうものも知っている人がつけてくれるわけです。

あるいは自分なりのまとめ方をつくる。私のコレクションをつくるなんていうことも可能にわけです。それはあくまでうまい、みんなが共有できる形で公開されているというのが前提だと。やっぱり別のアーカイブにおいても、こういう利用者の視点というのを考えたやり方というのが必要かなということを、今日お伝えしたかったということで、ちょっと時間がかかりましたが以上です。

【杉本座長】 どうもありがとうございました。いろんなものをつないでいくというところの重要性についてお話しいただいたかと思います。そうしたら、何かクイックエッセイはございますでしょうか。

もしよろしければ、続いて加茂構成員にお話をお願いしたいと思います。最初の新さんのお話がクラウドで、コンテンツを入れて共有するということで、また今、武田先生のお話が、アクセスを利用するほうで共有するという話で、ちょうどこれから加茂構成員にお話いただくのは、その共有されるハイエンドコンテンツのお話かなと理解しております。

【加茂委員】 じゃ、よろしいでしょうか。

【杉本座長】 お願いします。

【加茂委員】 こんにちは。加茂でございます。よろしく申し上げます。

今、杉本先生からおっしゃっていただいたように、私のほうは少しコンテンツの話をしたそうと思って準備してきました。私は実際に凸版印刷のほうで担当している分野の話をしように思っていますので、会社全体でこういうことをやっているという話ではなくて、我々の分野ではこういうことをやっているというご紹介に少しなると思います。そういう意味ではこちらで議論されているデータベースとは、だんだんと距離が遠い話かと思うんですけども、でも我々としては、こういうつくったものがそういうところにつながっていければなと思っておりますので、ちょっと距離感はあると思いますが聞いていただければ幸いです。

私は印刷会社にずっと長くいるんですけども、印刷会社の基盤技術というのは、1つの版から大量に印刷物をつくるということはもちろんなんですけれども、その中でテレビ

とか、そういったメディアと少し違うところは、印刷というのは出力サイズが非常にまちまちでして、小さな印刷物も刷りますし、壁一面大きな印刷物をつくったりもするんです。

ですので、20年ぐらい前からどんどん印刷会社の、我々の会社の中でもデジタル化が、画像も文字も進んできたんですけども、そういう中でもそういう特徴がありまして、一番大きいサイズを出力できるような、あるいは計算できるような仕組みが最初からないと思ってこられないという意味があって、言いかえれば、高精細な画像処理をするということが、いつもDNAのようにしみついているというところがあります。

もう一つは、印刷というのはお客様の大事な商品を扱うケースが非常に多いものですから、非常にシビアな色を、長い印刷の工程の中を経て、最後に、紙も違う、インクも違うというようなところで、なるべく正確に出したい色を出すということが必要で、そういう意味では研究所の中に色を専門に研究するチームが長年育っておりまして、我々はカラーマネジメントと言っていますけれども、そういった正確な色を出すというのがちょっとしみついているというところがあります。

デジタルアーカイブという仕事も、我々はそういう非常に大きな画像が扱えるシステムが大分昔から入っていたものですから、そういう意味ではデジタルアーカイブという言葉が90年代中ごろに出る以前から、そういう仕事を実は行ってきました。その中でも今申し上げたような高精細であるとか、色を正確に持ってこようという努力は続けてきたんですけども、最近では三次元の計測なんていうのも大分行われるようになってきました。ちょっと今日はその話をさせていただきたいと思っております。

文化財をデジタル化して、今申し上げたような技術を使って情報化する、デジタル化するということをした後に、これはいろんな分野で活用はされていくのかなというところで、もちろん1つは保存をするということもさることながら、最近ではもうデジタル情報を使って研究そのものを進める、それを使って公開するというようなことだと思います。

デジタル情報というのは、今議論が非常にされていますけれども、保存について非常に可能性が高いというようなお話もあります。もう一つ、デジタルだからこそというデジタル表現の技術を使って、いろんな文化財の側面を公開していくという利点もあるかと思えます。その辺をちょっとスポットライトを当てて、我々は仕事をしております。

印刷会社ですので、こちらはグーテンベルクの42行聖書ですけど、このバチカンの教皇図書館のアーカイブの仕事をしていたときに、このグーテンベルクもやったんですけど、そういう平面のものが最初は多かったんです。右のほうはフィッツィーですけども、こ

ういう絵画とか文書とか、それから古地図とか、わりと平面のものをデータ化する、90年代のもう少し前、80年代ぐらいからそういう仕事にかかわってきました。国内のものだけではなくて、海外のものも含めてやってきました。

ここしばらく立体物もデータ化をしようという、我々自身の研究なりお仕事なりがだんだん増えてきて、最初は、これは田良島先生にもいろいろご協力いただいてつくってきたものですが、こういう小さな八橋蒔絵のようなものとか、それからだんだん大きくなって建造物とか。最近では古墳なんというのを、毎年九州の彩色古墳をテーマにして計測をしております。

だんだん対象物が大きくなってきて、三次元のものも増えてきているとなると、こういうものをどうやって公開するかということで、もちろん我々の中にも映像のチームもありますので、DVDをつくったり、そういう映像のコンテンツをつくるということもやるんですけども、私どもは研究所と一体になっている組織ですので、バーチャルリアリティという手法を一つ取り上げて、14年ほど前から行っているということです。今日はそのあたりのお話をさせていただきたいと思っております。

細かいということなんですけれども、我々も仕事でやるケースが多いものですから、やみくもに細かければよいと思って細かくしているわけではなくて、これは対象物がどういう対象物であって、対象物の文化財のどういうところが将来的に知る必要が出てくるかというのは、この情報のもとの対象物の素質というところから割り出すのと、もう一つは、それをどういうデバイスで、どのぐらいの解像度のデバイスで出力して、どこにスポットライトを当てて拡大したり、見る必要があるのかということ、最初に想定します。それを逆算させていって、カメラを決め、精細度を決めということでやっていくようにしています。

ですので、結果的に細かいという一言になってしまうんですけども、それは目的があってしているということになる。もちろんですからインデックスのように、それが何であるかということがわかる情報を網羅的にとることを、別にここでは否定しているわけではありません。そういう仕事はそういう仕事でもあるんですけども、ここではそんなふうにして撮っています。ですからここも対象物と出口の中で仕様が決まってくるということにしてあります。

これは色を計測しているところです。これは東京文化財研究所の委託として我々の研究者と一緒に、アジャンタの遺跡の壁面の分光特性をはかっているところですけれど

も、こんなことをやります。

特に最近デジタルカメラも我々は普通に対応するんですが、カメラもなかなか基本的にはローデータで撮って持ってくるんですけれども、レンズとか、それからカメラのCCDなりの特性とか、エンジンの特性とかによって、カメラもそれぞれ癖があるというところで、本物の色は何ですかというそのところを、単にRGBの信号値からでは割り出せないで、正確な対象物の色をはかっておくということを同時並行にやります。もちろんカメラのプロファイルも保存しながら、最終の出力物の精度を高めていくということをやっています。

これは三次元計測をしているところで、阿修羅さんを撮らせていただいたところですが、こんなふうに形状データをとったりもしています。必ずしも形状データをとれるものでないときには、修復図面とかいったものでデータ化をしていくケースもあります。

我々はこういう文化財をデジタル化するという仕事を、80年代ぐらいからいろいろやってきたんですけれども、最初のうちは、文化財は本物があれば十分なんだとおっしゃる先生方も多数いらっやって、こんなことしなくてもいいんだよということをおっしゃっていた方もたくさんいらっやいます。

それはそれで理解できますし、文化財は当然かけがえのないものだと思っているんですけれども、そうはいってもここで議論されているように、情報をアーカイブしておいて、それを将来にわたって活用するということは、今の時代だんだん普通になってきて、あまりそういう議論がなくなったかとは思いますが、我々が目指しているのは、あくまでも文化財に超えて情報がある、そういうことを目指しているのではなくて、もちろん物そのものの物理的な特性もさることながら、研究者の先生方の情報をもとにして、素材としての情報から、もう少し価値をつけていって、文化財と補完関係が保てるような、新しい価値の創造ができてくると、ごらんになる方ももう少しまた、文化財に親しみを持って入ってくれる、そういったこともあるのかなと思いつくっています。

こうやって三次元の計測をしたり、高精細な画像を撮ったりするんですけれども、コンテンツの制作過程においては、学術的な監修の時間をかなり多く割きます。大体こういうものをつくろうと思うと半年から1年ぐらい必要なもので、時間をかけるんですけれども、半分ぐらいはこういう取材にかかっていくんです。こんなことをしてデータ化していきます。

結果としてこういうバーチャルリアリティというもので見せるのが、我々のところの仕

事なんですけれども、今ごらんになっていただいているのは、凸版にある小石川のVRのシアターです。奥の壁面に映像が流れるんですけれども、基本的には人が立って、お客さんの前でとうとうと暗記したナレーションをお話する、ナレーションのかわりにしゃべるという意味ではなくて、お客さんの顔色を見ながら、いざなっていくというような心がけでご案内していくつもりで、ナビゲーターと言っていますけれども、もちろんオートマチックに、シーケンシャルに起承転結のシナリオをつくってお見せするというのも、それはそれでやってもいるんですけれども、できれば横道にそれながらご案内することができるようというつもりで、この映像はリアルタイムのレンダリングをして、文化財の中を散策するように、一緒に歩くような感覚をつくっていかうと考えてやっています。

これは機材ですけど、小石川のシアターはプロジェクター3面で、ブレンディングして、ちょっと幅広目にシリンダーのようなスクリーンに出しているんです。これはこれでお客さんを包み込むような雰囲気、臨場感を出すために、こうしているんですけれども、我々はハードメーカーでないので、これはその都度合わせていけばいいかなと思っておりまして、たまたま我々はこんなになっている。

お客さんに対して人がご案内していくことをすると、何が生まれるかという、お客さんの反応がいろいろ我々にとっては楽しみなんです。ナビゲーターの人が、お客さん、鑑賞者に対して一緒に文化財の情報なんかをご案内していくと、お客さんのほうから徐々に反応が生まれて、ここはどうなのよとかいう質問なり反応が出てくる。そういうことによって、またシナリオを少しずつ考えながら。勝手に変えながらではないんです。これは監修者の方に確認をしながら、伝えていい情報を伝えていくということをやっています。ですから我々はこういうものをやっていくと、1つの対象の情報からたくさんのコンテンツが年を追うごとに増えていくということがあります。実際にそうしています。

事例のことを少しお話ししますが、これは昨年、遷都1300年の事業のときに我々が作ったコンテンツです。これは平城京全域を、奈文研の先生方に入ってください可視化したものです。ちょっと細かいですけども、民家がかちゃかちゃとあります。これも2万5,000軒をもうちょっと超えるぐらいの民家を一一つつつくって行って、まち全体をつくっていった。この中を説明しながら案内するというもの。

これはポンペイのコンテンツです。今のポンペイの遺跡は左上のようにこんなになっていますよ、昔はこうだったんですよ。この中に入っていくと皆さんが食事をしていたようなところとか、壁画が復元されて、データ上では入っているというようなことです。

それから、なかなか中に入っていけない遺跡というのが案外多くて、特に我々が毎年やっている古墳なんかはそうなんですけれども、古墳自体もどんどん傷みが進んでいく、色もあせていく状況があって、その中で公開ができないということもあるので、この辺は今、正確に色と形の情報を毎年はかって、右のほうは九州の古墳ですけれども、これは東京大学の池内研究室と共同研究で、もう6年になるんですけれども、古墳の対象物に合わせて計測手法を研究し、それができたところで計測に行くということを行っています。これはコンテンツにして、九州国立博物館で公開を続けています。

左のほうはマヤの遺跡の地下にある神殿の一部ですけれども、こんなのもやっています。

あと、これは故宮の紫禁城なんですけれども、2000年から故宮——中国北京の故宮ですけれども、凸版印刷で共同研究を始めまして、すべての建物をデータ化をしようということです。実は故宮の中でも情報が全部整っているわけではないので、建造物の情報も文物の情報を研究しながら、それをデータ化して可視化していくということをずっと継続していきまして、どんどん情報が積み重なっていっています。これは故宮の敷地内に凸版との共同研究所を建てていきまして、ここにも研究者用のシアターがあります。

シアターという意味では、これは東京国立博物館さんとの共同の事業として、実験的ではありますがけれども、毎週金土日公開していきまして、博物館というのは物が収蔵されていて公開されているところ、もちろんそこには研究の成果がある。それらを足して、そのものがどう使われていたのか、どこにあるのか、そういう臨場感を加えて、博物館の公開とプラスするような補完関係をつくりながら、こういう公開をやっております。

海外での博物館ですと、我々がかかっているところだと、左上が故宮の研究所です。これは古い建物のデザインですけれども、2003年でしたか、新しく建てた建物です。ちょっと囲うような建物になっていますけれども、三次元計測をしたり、もちろん文化財そのものの研究をしたり、それから研究者に対して公開するというような施設があります。上の真ん中がその故宮の中のシアターです。調印式なんかがあります。

あとはホンジュラスの国立博物館ですとか、そんなところに我々がかかわった公開の施設があります。

最近では、博物館で行われる大型の企画展にあわせてコンテンツをつくる機会が多くなってきていまして、「奈良の大仏の世界」という、2010年にやりましたけれども、こういった大型の企画展。大仏さんにまつわる品々は持ってこられることができますけれども、大仏さんを持ってくることはできないので、これは大仏さんの手のひらを原寸大で表示し

てみたりとか、そういうことをやりながら、企画展全体の臨場感を高めていくということ
を目的にしまして、こういうコンテンツをつくりました。

こんなようなコンテンツを企画展とあわせるようになってから、実はごらんになって
ただお客さんの数も大分増えてきて、我々はずっと積算しているんですけども、
今年の冬の段階で700万人を超えるお客さんが、実は我々のこういうコンテンツをごら
んになっていただいています。このコンテンツは同時はかなり精細度が高いので、あるシ
ーンを撮って印刷物にしたり、それからこういう公開中にDVDにして販売したりします。
格安に販売をしているつもりなんですけれども、結構見て買っていただく。そのものずば
りではなくて、プラスアルファのコンテンツを入れてDVDにしているんですけども、
結構売れ行きもいいと思います。

ちょっとかいつまんで持ってきましたのですけれども、今のはクレムリン宮殿です。

これはもちろん兵馬俑。兵馬俑は彩色された状態まで復元するようなコンテンツになっ
ています。

これは金閣寺です。パリで公開をしました。これは江戸城の本丸の中です。こんなふう
にして、たくさんの方の研究者の人たちとやりますし、すごくたくさん素材の写真も撮ります。
計測したり、修復情報としてのデータ、形のデータからこういうものです。これは色をは
かっているところです。

これはアンコールワット、バイヨンです。これも東大の池内先生が、もう何年もかかっ
てはかっている膨大なデータがありまして、凸版の高速のレンダリングのプログラムを使
って、リアルタイムレンダリングで、自由にこの細々したところは入れるようになってい
ます。

これは江戸城です。これはマヤの遺跡ですね。地下に埋葬されている神殿です。アンコ
ールワットはほんとうに膨大なデータです。

これはナスカの地上絵です。上空から見るとなかなか見えにくいところを、こういうC
Gで少し見やすくしているんですが、おりていくと石の一つ一つが寄せられているとい
うところまでわかる、どうできているかというのがわかるような解像度も出しています。
こんなふうにお客様を情報と引き寄せながら、体温を近づけながら案内していくとい
うようなやり方をしています。

これは東博さんの。企画展ですとこういうふうにお客さんが入っている。

これは今もつくり続けている北京です。一番最初につくったコンテンツで。これは違い

ますけど。ちょっと早回しで。

ということで、事例ばかりになって申しわけないんですが、実はこういうコンテンツを1個つくろうと思いますと、普通程度のコンテンツの規模のものでも、さっきちょっと写真撮っているシーンがありましたけれども、我々の写真の撮り方ってちょっと変わってまして、柱の一つ一つを撮っていったり、床のタイルまで撮っていったり、屋根のかわらを全部撮っていったりとか、そういう撮り方をしていくんです。ですので、記録としても現状の文化財の状況が非常にわかるような素材がかなり多く残ります。1週間ぐらい取材をやりますと、1万カットとか、場合によっては2万カットぐらいデータが撮れるんです。

それがあって、それからコンテンツができてくるということになりますので、こういうものもできれば、プラットフォームといいますか、データベースがきちっと整備されていけば、そういうところにひもづかせていただくと、文化財の情報って非常に息が長いので、いろんな活用もまた将来増えてくるのかなと思って、こういうところに期待して出させていただいております。

私のほうからの報告は以上とさせていただきます。ありがとうございます。

【杉本座長】 どうもありがとうございました。

いかがでしょうか、何か今の加茂構成員からのご発表に関しても、あるいはこれまでの新構成員、それから武田構成員からのご発表に関しても、ちょっとここで5分ぐらい。はい、宮澤先生、お願いします。

【宮澤委員】 新さんへのご質問なんですけれども、この質問は要するに、長期保存に関するコストに関係しての質問です。仮に今何かやっているようなデジタルアーカイブを組織内でやっていたとして、それを例えばP a a Sでクラウドに移すというプロジェクトをやったとします。それに関する費用と、それから既にP a a Sでやっている、クラウドでやっているものを別のクラウドに乗りかえる、そういうことをやろうとしたときに、どの程度の差になりますか。同じぐらいかかりますか、それとも安くなりますか。

【新委員】 大変難しい質問なんですけれども、そういう意味では課金体系自体も事業者によって大きく違いますので、ストレージはお安いんですが、ネットワークの体系を使うところに課金がかかるというスタイルのこともありますし、逆にネットワークは使えますけれども、ストレージの移行にかかりますというパターンがありますので、これは具体例がないと、ちょっと今お答えできないです。

【宮澤委員】　　そういうお金の部分を除いて、仮にそういうプロジェクトをやるとしたら、どれぐらいの人間がどれぐらいかかるかという点で考えるとどうでしょうか。

【新委員】　　申しわけないです。今ここで即答はできかねますが、どのぐらいのイメージで。

【宮澤委員】　　私が一番知りたいのは、同じぐらいかかるものか、確実に楽になるものなのか、少しぐらいは楽になるかなというものなのか、その程度でもよろしいです。

【新委員】　　正直、一応 I I J としてのクラウドの事例もいろいろありますけれども、全くケース・バイ・ケースとしか申し上げられないんですが、クラウドに移行される場合というのは、移行だけではなくて、その後の管理のコストを一緒に考えていただくというのが多分一つ、大きい点になりまして、これはほんとうに私の私見なんですけれども、管理というのは人がかかりますので、人を人材として確保することと、システムに移行することというところで、てんびんにかけていただくような考え方になると思いますので、事業者のほうは、今それが安いですよとして実際売っているわけですけれども、それはいろいろな状況によって、その点を含んで考えていただくというのが。ちょっと玉虫色の答えで申しわけないです。一般には人よりはシステムのほうがお安く見積もられるかなとは思っています。

【宮澤委員】　　では、クラウドと同じ P a a S で乗りかえるとすれば、その管理のための人というのは、基本的には同じことをやるので済むでしょうか。

【新委員】　　それは、もともと今クラウドにあるものを違うところに移行する場合でしょうか。

【宮澤委員】　　はい、その場合です。同じ P a a S の形だったとしてです。

【新委員】　　同じ P a a S の形であれば、基本的には同じになると思いますが、その P a a S が全く同じかどうかを判断するというのは、また別のところでやる必要があります。

【宮澤委員】　　どうも。

【杉本座長】　　ちょっとクラウドの話で私もお伺いしたいんですけども、いわゆる S a a S、Software as a Service のところで、例えば今 Google ですとか、そうしたところでのドキュメントだとかメールなどのサービスというのは一般的かと思うんですけども、例えば電子政府向け S a a S だとか、あるいは美術館向け、博物館向け、公文書館向け S a a S という、ある種クラウドって、レイヤーの構造をうまいこととっていくというものかなと思うんですが、ある種の決まったコミュニティ向けのそういう S a a S という

アクティビティーというものはあるでしょうか。

例えば、どこかの国でe-Government用のクラウドのサービスのレイヤーというのをつくっていますとか。

【新委員】 それは国が用意されるのか、事業者がやっているということでしょうか。

【杉本座長】 国か、あるいはそういう国向けにやっているか。

【新委員】 海外について実はあまり知らないんですけども、電子政府向けというのは、多分むしろ総務省さん自身でやられていると思うんです。それ以外の実験レベルで事業的なものとして、例えば医療クラウドですとか農業クラウドという形で、ある程度ユーザーの方を特化した形の基盤というのは、今実験レベルプラス少し実証を行っているはずなんです。

【杉本座長】 そうしますと、例えばそういうところでいくと、標準化が多分大きな問題になるかと思うんですけども、その標準化のアクティビティーで、こういう領域であれば何年ぐらい先には標準化が進んで使えるようになりますよと。例えばMLA関係だと、そういうアクティビティーとしてこういうのがあって、あるかどうか問題ですけども、ほかのところ例えば例があると、何年ぐらいたつとこういうのができますでしょうねというのがあれば、横目で見ても動くこともあるでしょうから、そういう事例をご存じでしたら教えていただきたいんですが。

【新委員】 残念ながら具体例は私も知らないんです。

【阪本審議官】 座長がおっしゃるような形で、今どちらかというともまだ取っかかりなんですけれども、分野ごとに多分それぞれ特性があるだろうということで、その分野の特性は何なのかということ、それぞれ皆さんで議論していただいているというような状況にありますので、電子政府で申しますと、霞が関クラウドのアプローチとか、自治体クラウドのアプローチ、そのあたりは具体的なベースとしてある。多分また医療のほうは医療で、どういうクラウドの展開ができるのかということがありますので、それぞれの分野ごとと、多分分野を横断する共通部分と当然あると思うので、今それをそれぞれいろんなパターンでいろいろ試行錯誤しているという状況でございます。

多分それがある程度固まったら、標準化とかガイドラインとか、できれば我々もそちらのほうに少し貢献できるような形で展開できればなと思っています。

【杉本座長】 ありがとうございます。きっと大きいところもそうですけど、小さいところからすると、ほんとうに頼りにしたいサービスかなと思います。田良島さん、どうぞ。

【田良島委員】 博物館、美術館向けのSaaSですと、例えば具体的に名前を挙げちゃうと、凸版印刷さんのMuseScopeですとか、それから廣済堂さんのミュージアム・データバンクですとか、そのクラウドを利用した個別の博物館向けのサービスというのは、ぼちぼちと出ているところかなと思います。

ただ一番の難点は、武田先生がご指摘になりましたように、メタデータが非常にばらばらなので、そういうまとめるメリットがなかなかない、難しいというところだろうと思います。

【杉本座長】 ということは、要はコンテンツそのものをクラウド化しないといけないということですね。

【田良島委員】 ということですね。あと、文化遺産オンラインもそういう方向に何となくなくなって、しょうかという感じもないでもないですね。

【杉本座長】 ほか、いかがでしょうか。今この時点でのご質問、あるいはご意見。

【宮澤委員】 もう一つよろしいですか。

【杉本座長】 はい。

【宮澤委員】 コストのことばかりで何なんですけど、加茂さんへのご質問なんですけど、要するに単なる二次元の画像のアーカイブのシステムをつくるのと、このVRのできるようなデータをつくるのと、どの程度のコストというか、手間の差があるものでしょうか。

【加茂委員】 これも申しわけございません、一概にということとは言えないんです、条件によるんです。

【宮澤委員】 もちろんそれはわかります。

【加茂委員】 やっぱり手間も時間もかかりますので、その分コストはどうしてもはね上がってしまいますね、現実問題は。ですので、ああいう今日お見せしたようなものをつくるときには、やっぱりある程度の条件が整わないと難しいと思っています。現状、具体的な金額はここではちょっといろいろ申し上げられないんですけれども、例えば企画展なんかのときには、企画展の主催者と、半々じゃない場合もありますけれども持ち寄ってつくって、それを回収していくというシナリオを書きます。

企業の側からあれですけれども、景気のいい時代ですと、どこもやはりもうかり過ぎちゃったお金で、どんどんこういう文化事業に貢献してみたいなところで成立していたところも、現実はあると思いますけれども、これからはそうではないということと、そうであ

ってはいけなくて、金額の多少にかかわらず、いろんな意味できちっと回収ができていて、それで次のコンテンツを生む仕組みを考えていかないと。システムに張りついたようなコンテンツが1個あって、10年先に行っても同じものが流れているということが起きてはいけないと思っています。

ですから、おっしゃる質問のまさにそこがものすごく大事で、そのマネジメントをきちっと考えていかないと、いつまでたっても活用のところは、必ずしも何千万もかかるものだけの話ではなくて、文化財の対象が量が多いですから、少なくとも10万円でも100万円でもかかってコンテンツをつくっても、それをつくりっ放しでどんどんこれからもいけるかという、それはそうではないと思うんです。

ですから我々のようなコンテンツが海外に行って博物館で見ると、別料金を取って上映している場合もかなり多いです。これは我々調査もしましたけれども、場合によっては、その博物館の入場料金よりもはるかに高い金額で上映しているというケースが結構多くて、それでもお客さんが入っているというケースはありますし、そこはやっぱりすごく大事なところだと。我々は企業なので、特にそう思うのかもしれませんが。

【宮澤委員】 すいません。アイデアをつかむためにですが、仮に1部屋分のふすま絵を二次元の画像だけ撮るというのと、バーチャルリアリティ可能なような形でセットを撮るというのと、100倍かかるものか、1,000倍かかるものか、10倍のものか。

【加茂委員】 100倍かかるということはあり得ないですし、それは仕事として成立しないので、ちょっと10倍ぐらいかかっちゃう場合もありますし。ただ、ほんとうにこれは申し上げにくいんですけども、情報があるものからスタートするとわりと楽なんです。例えば修復図面が過去にあって、部屋のきちとした図面を持っていれば、そこからスタートすれば、別に計測したりするコストはなくなりますし。

ですので、ちょっとこれは方便かもしれませんが、こういう情報がきちっと蓄積されている場所にどんどん行って、そこに我々がたどりつけることが簡単にできるのであれば、どんどん制作のコストは下げることができると思います。それが分散していたりすると、なかなかこれも大変で。ですので、制作時間の半分ぐらいを調査にかけたり、それから許諾を取るという手間もものすごくあります。

単価が決まっているフィルムを借りてくるのであれば、それは簡単なんですけれども、文化財ホルダーの中には、なかなかそれだけでは交渉が済まないものもあります。どう使うのかという解釈をきちっとしていただいた上でないと、お金の問題だけじゃ片づかない

問題もありますから、そういうところに時間がかかる。それもすべてコストですので。

【宮澤委員】 ありがとうございます。

【加茂委員】 すいません、お金の話をきちっとここでリストを出せばいいんですけど、それはちょっといろいろ。

【杉本座長】 どうもありがとうございました。

そうしたらちょっと時間の都合もありまして、次のほうに進んでいきたいと思いますが、どうもありがとうございます。後でもし時間があれば、また戻って議論できるかと思えます。

それでは次、事務局からのお話でお願いいたしたいと思えます。

【松田情報流通振興課統括補佐】 お手元の資料5-4につきましてご説明申し上げたいと思えます。

なお、今回のこの研究会の全体のスケジュールでございますけれども、これまで2月から開始してまいりまして、今回で第5回ということでございます。さまざまな皆様からプレゼンテーションを、また意見交換もいただいておりますので、これらをもとに次回、おそらく10月ぐらいになるかと思うんですが、研究会としての提言の取りまとめに向けて、何らかのたたき台が示せればと思っておりますので、またメーリングリストなどで、この論点について足りないんじゃないかとか、この点について意見を出すとか、ぜひご協力いただければと思っておりますので、どうぞよろしくお願いいたします。

それでは資料5-4についてご説明申し上げます。5-4は、本年度の調査研究の資料でございますので、1つ目が実証実験、2つ目が調査の関係、3つ目がつくるべきガイドラインのイメージでございます。5-4-3は席上配布とさせていただきます。

まず、5-4-1でございますけれども、お聞きいただきまして、スライド番号2、3あたりをごらんいただければと思うんですが、デジタルアーカイブについて皆様にご議論いただいておりますので、さまざまに論点があるかと思うんですが、ひとまずこちらでは、小規模な図書館、博物館、美術館などにつきまして、デジタルアーカイブが十分進んでいないのではないか、そもそも構築がされていないという現状を踏まえて、そういった小規模な記録機関におきましてもデジタルアーカイブを構築していけるよう、あるいは大きな機関とも連携をしていけるよう、そのための実証実験が行えればと考えております。

スライドの3にございますように、博物館・美術館、図書館、文書館の持つ資源をデジタルアーカイブにて構築して、相互に連携して利便性を高めよう、そういった実験でござ

います。その中で今回は、都道府県立の図書館が、一つの中核的な支援機関として役割を担う、そういったモデルを実際に実証してみてもどうかと考えております。

スライドの4ページ、あるいは5ページにございますけれども、これはイメージ図でございます。

若干わかりにくいので、スライドの7ページをごらんいただければと思います。右側に中央機関とありますけれども、今回はこちらの構成員でもあります国立国会図書館を中央機関といたしまして、真ん中に県立図書館がございますが、山崎構成員のところの秋田県立図書館、また、参加機関として秋田文学資料館、秋田県立近代美術館、あるいは秋田県の公文書館、こういったところにご参画いただきまして、それぞれ参加機関については所蔵目録の整理、あるいはメタデータの登録を行う、また、秋田県立図書館にこれを支援いただくというようなイメージを考えております。

さらにそこでできた目録情報、メタデータを、平成22年度の総務省の事業で構築いたしました、メタデータ情報基盤、MetaBridgeがございますけれども、こちらに登録をする。これによってさらに国会図書館のNDLサーチ、あるいはPORTAへの連携を図ることによりまして、都道府県立の図書館が中間集約を行い、県内の参加機関の支援を行うような、そういったモデルの実証が行えればと思っております。

スライド番号5番に戻っていただきますと、本年度はこういった形で進めることによりまして、人間系の手続や、あるいは作業を中心に実証実験を行いまして、運用していく上での課題の洗い出し、ノウハウの収集を行いまして、ガイドライン等に生かしていければと考えております。

さてもう一つ、資料5-4-2でございます。今回いろいろテーマがある中で、メーリングリストでも話題になっておりますデジタルアーカイブの長期の利用、あるいは保存をどう保証していこうか、この点について調査を行おうと考えております。長期利用するために有用な技術、あるいは管理手法や実例に基づく成功、失敗事例をまとめまして、これからデジタルアーカイブを構築する方々への参考となる調査にしていきたいと考えております。

おめくりいただいて、2ページにあるようなヒアリング先候補からのヒアリングを行います。また、青空文庫さんだとか、グッティイメージズさんだとか、そういうところにもお願いできればと思っております、4ページにありますようなヒアリング項目、特にどうという失敗事例があったかというようなことを聞いていければと考えております。また、

海外の調査も行おうと考えております。

以上が長期保存に関する調査の概要でございます。

最後に席上配布させていただいております、21年度までのガイドラインの素案が各種できておりますけれども、これを統合いたしまして、「デジタルアーカイブの構築・連携のためのガイドライン（たたき台）」としております。これについても今後、いろいろご意見をいただきながら、完成させていただきたいと考えております。

以上でございます。

【杉本座長】 どうもありがとうございました。そうしましたらば、今お話しいただきましたことに関してのご意見、ご質問などお願いいたしたいと思っております。

【入江委員】 お願いなんですけど、データの長期保存の話と具体的な作業のところを、どこかで合わせていただきたいというか、例えば先ほどのクラウドの話でいくと、ウェブに公開するようなデータのクラウド化というのは別に難しくないと思うんですけど、そのバックにあるでかいデータを持っているので、これをどうするというのが問題で、クラウドに乗っけるとしても、乗っけても後で何も加工できなくなっちゃうので。

多分具体的に、そういう小規模の図書館、博物館の方々が電子化したときに、もうウェブに乗っけるのはそれでいいとしても、撮ったものそのものがあるわけですから、それをどう保存するのかという議論は、やっぱりどこかであるわけですね。だからその議論とそのデータの長期保存というものを、具体的なワークフローとして議論できるような形にまとめていただければなど、ちょっと思います。

【杉本座長】 どうもありがとうございます。どうぞ自由にご意見をお出しいただければと思います。はい、どうぞ。

【武田座長代理】 実験のほうは対象は公文書館等とか、どの辺の、いわゆる対象物は何を想定しているんでしょうか。

【松田情報流通振興課統括補佐】 お手元の資料には実はまだ明記していないところなんですけれども、秋田県立図書館さんに中核になっていただいて、秋田文学資料館さん、それから秋田の近代美術館さん、秋田の公文書館さんですので、少なくとも公文書、美術品、あるいは文学資料館ですので、地域の文学資料ということになるのかと思います。

もしよろしければ山崎さん、補足をお願いいたします。

【山崎委員】 私が関係していることなので、じゃ、私のほうから。

基本的には目録データです。エクセルで数千件レベルで各館で持っていますので、さま

ざまな形態で、要するに項目がばらばらなんです、当然のことながら。それから、同じ館の中でもばらばらになっているというケースを幾つかピックアップしてもらって。例えば文学資料館なども博物資料もありますし、当然ながら初刊とか1点資料、それからもちろん本もありますので、既に文学資料館の中でも項目がばらばらになっているんです。今館内でも検討しているようですけども、そういうものを1回提供していただいて、データベースの中で統合を図っていくということを考えています。

美術館さんも同じです。近代美術館さんの場合にはウェブに公開していないんですけども、内部的に整理しているものがおありだということで、それを提供していただくと。公文書館さんはウェブ上に、エクセルのファイルは提供しているけれど、検索ができないので、そういう意味でそれを検索できるような形での提供を望んでいらっしゃるということで、説得できました。

ただ博物館さんに関しては、この趣旨がよくわからないということで参加いただけなかったんです。私も直接行ってお話したわけではないので、電話で話した程度なので、うまく説明できなかったんですけど、何が生まれるか、やはりこういうメリットというものを書いて実験で相当明確にしないと。要するに参加していただくのは、何がこれによって得られるかというものをなかなか示さないと、難しいなとは一応感じました。

ちょうど武田先生が美術館のサイトを今お話しされたので、私、ちょっと聞こうかと思ってはいたんですけども、15館の美術館が参加されたときに、要するにどういう反応だったのかとか、あるいはどういうそのときの説得の課題があったのかなというのも、あわせて逆に質問を。

【武田座長代理】 あれはウェブにあるものを、我々が勝手にスクレイピングしているんです。

【山崎委員】 勝手にやってきたんですか。

【武田座長代理】 横浜美術館を除いて許可を得ていません。横浜美術館は実は、ちょっと横浜のローカルデータを使うとかいうので、財団さんとお話をしたりして。

【山崎委員】 それで近代美術館、秋田県で22データというのは、私は逆に言うとおかしいなと思ったんだけど、けたが私の頼んでいるデータとはちょっと違っていたので、それを。個人的にお話ししていくと、メタデータの説明から始めなければいけないんですね、特に博物館、美術館の世界の方々にお話しするときは。公文書館もある部分そうだったんですけど、メタデータって何という話からしなきゃいけなかったんで、もう数時間か

けてお願いしないと、なかなかその了解を得られないというのが実際のところにあります。

要するに手前でとまっているんです。エクセルもつくった、デジタルデータもつくったけれど、それをどうしたらいいかもわからないでいるというのが、多分地方のこういう機関の現状じゃないのかなと思います。

ただ、この会議を見ている方は結構いらっしゃるので、それはかなり効果が。私は何でこの会議をやっているのか、最初はよくわからなかったんですけど。今私のいるところに、参加している中でいろいろ議事録が出ていますが、あの議事録を見て声をかけてくれた企業とか、あるいは機関があるので、このガイドラインを作成すれば、意見ではないんですけど、感想としてかなり私は効果があるのかなと思って、最近やっと何か意味がわかってきたという感じがしました。余計なことです。

【杉本座長】 どうもありがとうございます。じゃ、松田さん。

【松田情報流通振興課統括補佐】 山崎構成員がおっしゃるとおりでございまして、この資料5-4-1の3ページ目にも、「参加いただく機関へのメリット」ということで、県内他館のノウハウを活用しながら、自館のメタデータの整理・蓄積ができるだとか、眠っている地域の文化財資源の整理を行うことができるだとか、いろんな資料保存機関に、連携しやすいデジタルアーカイブを作るといいよという投げかけを行うための言葉を書いているんですけども、キラーとなるような、必要性を実感頂くアプローチというのはなかなか確立されていない。

今回こういう実証実験を通じて、成功モデルじゃないですけども、非常に意味があることなんだということをご理解いただき、そういう社会実験になればいいかなと思っておりまして、ぜひ山崎構成員にお力添えいただければと考えております。よろしく願いいたします。

【杉本座長】 安達先生。

【安達委員】 実証実験を行う上で、博物館からの協力が得られないということで、それは本質的なことを意味しているんだと思うんですが、技術的な実証実験という意味で、やはり物を扱っている機関、博物館あるいは資料館の協力、それを対象にしないと、本来の確認はある意味できないかなという気がしております。

というのは、図書館、美術館、文書館、いずれも扱う資料は大抵名前があって、作者があるんですが、作者等がない物、資料まで含めてどう共通に扱うことができるかというこ

との技術的な意味の検証をするとすると、やっぱり物を扱う博物館を何とか理解を得ていただいて検証の対象にするということが、求められるかなと思います。

【杉本座長】 ありがとうございます。いかがでしょうか。

そういう意味では、保存という視点からいきますと、例えば国立国会図書館でも随分長くデジタル化を進めてこられていますし、それから先ほど加茂さんからのお話で、最初は80年代から始まっていますよというお話があったので、ある意味で10年あるいは20年を超えるデジタルコンテンツを扱ってこられた経験というのは、既にお持ちかなと思うんです。

例えば加茂さんのところでやってこられた非常にハイエンドなものというのと、特にまた技術的に進歩が激しいところでしょうし、例えばそういうところで、要はコンテンツの保存ということに関して、こういうところに問題がありましたよとか、あるいは今まで全然問題はありませんとか、そうしたことをもしご経験としてお持ちであれば、教えていただきたいと思うんです。

またそれと同じことで、国立国会図書館でもやっぱり長くやられてこられていますので、そこでのご経験をもしちょっと披露いただければ、これからの調査に役に立つかなと思いますので、お願いできますか。まず、加茂さん、お願いできますでしょうか。

【加茂委員】 我々もコンテンツをつくるときに、1つの心がけとして、素材はきっちりできるだけのことをしてとっておきましょうよねと。いいかげんな言い方ですけども。デジタルの表現技術とかというのは、デバイスにかかわったりするところというのは、去年できなかったことが今年できるようになったりするわけです。なので、そういう表現の領域が広がってくると、今までちょっと上手に伝えられなかった部分を、新しい技術を使えば伝わっていくとか、技術的なインタラクティブではなくて、人とのインタラクティブが、よりとれるようになってきたりということが出てきたりもするんです。

つまり、これをずっと長く使いましょうねという情報というかデータと、それからもうこれちょっと変わってもいいよね、変わっていくのは当然だよねと思っておくべきものと、我々の場合は2つあるのかなと思います。

それからあと、ここでの議論は先行している図書館とか、美術館、博物館をどうするか、まずスタートはそういうところだと思うんですけども、実は企業といいますか、企業といってもいろいろあると思うんですけども、例えば映像を扱っていらっしゃるところとかが持っていらっしゃる資料、情報というのは、権利が非常に複雑ではなくてシンプルだ

と、私は逆に思うんですけれども、そういう資料もものすごくたくさんお持ちのところがあるんです。

特に文化財は映像にかかわるものというのは、例えば放送局ととらえてしまうと、これはどれだけあるかということよりも、その局の中でもそういうかかわりを持って撮ってきた方というのは、ものすごい何千時間という映像情報を持っていたりするんです。対象物がもう既にないものなんていうのは、そういう資料が一次資料化しているもので、非常に貴重になってきているものが既にあると思いますので、ぜひ次のステップはそういうところも、我々みたいな企業がやっているようなところも、少し視野に入れておいていただいて、我々がそういうところと、先ほどもちょっと申し上げたんですけれども、こういう基盤ができたところとうまく素材なりコンテンツなり。コンテンツもまたアーカイブに入っていく、こういう流れがそういう情報の根っこにつながるようなことを考えていただくと、結構膨らむんじゃないかなと思います。

あと、企業資料館とかあいつたところも、この日本の国を支えてきた非常に貴重な基盤のところの歴史が残っているということも、たくさんあると思いますので、海外で言えばそういう博物館も立派なものがありますし。ちょっとそんなふうに思いました。

【杉本座長】 例えばバーチャルリアリティの非常にハイエンドなコンテンツをつくられているわけですが、そのコンテンツそのものは、例えば20年たったときにはだめになっているかもしれないですね。要は技術が変わってしまっただけでもそのときであっても、例えばその素材というのが非常にきちっと残されていれば、それはそういう文化遺産の非常にハイクオリティーなデジタル化をしたコンテンツとしては、それでよしという感じでしょうか。それともやっぱりバーチャルリアリティの技術をつくってつくった作品というものも、一つのコンテンツとしてずっと持っていけないといけません。

【加茂委員】 いけないというよりも、残ってほしいとは思いますが、思いますけれども、そこはあまり固執しないほうがいいかなと思っていて、やっぱり受け取り側の感性も変わってくることもあるので、その時代のデジタルの表現技術に乗せかえていくということが出来るような仕組みとか、できる素材をちゃんととっておけばいいかなと思います。

例えば形のデータなんていうのはそれほど変わるわけじゃないですけど、表面質感を撮るカメラなんてどんどんよくなってきますし、高精細にこだわるわけではないですけども、粗いデータで撮るのも高精細で撮るのも手間は一緒ですから、コストを計算するとそ

んな変わらないんです。だからそういうものはきちっとっておきたい。直せばいいじゃないかと、ちょっと軽目に考えなきゃいけないところもあるんじゃないかと思っております。

【杉本座長】 ありがとうございます。

【武田座長代理】 ちなみにそのときに、その素材の部分というのは実際残しておくんですか。

【加茂委員】 我々が撮ったものは残しています。残しているというのは、データとしてももちろん残すということもさることながら、管理する責任があるんです。もちろん一つ一つ契約をとってやらせていただいていますけれども、これが散逸しない、変な使われ方をしないということも含めて管理しておく必要があるので、きっちり管理はしています。

【武田座長代理】 それも博物館等の保存するものとして、お返しするとか、そういう感じなんですか。

【加茂委員】 基本的には素材データは我々が持っています。ただ、それは個々の覚書なり契約なり、それが博物館さんにおさめるところ、場所があって、そうしたほうがいいものについては最初からお話し合いをすれば、そこにお預けしてしまうものもあります。現に、お寺さんなんかでやったもので、データを向こうに私どもと同じものをお渡ししておくというケースもあります。それは彼らがそれを将来必要と感じるから、その最初の話し合いでちゃんとしておきます。

【杉本座長】 どうもありがとうございました。すいません、大場さん、お願いできませんでしょうか。

【大場委員】 はい、国立国会図書館の大場です。国会図書館の経験ということなんですけれども、我々は紙の平面の本を画像でデジタル化することをやってきたということと、あともう一つは、パッケージ系電子出版物と言われるCD-ROMとかいったものを集めるという、その2面があります。

紙のほうに関しては、最初のうちは、まずフィルムで撮ってそれをデジタルにするという形でやっていきましたので、最悪フィルムが残っているので、紙の本直接ではなく、フィルムに戻れば再度デジタル化が可能であるという発想でやっていた面があったので、あまり大きな問題とはとらえていなかった側面は、最初のうちはあったと思います。

ところがデジタル化していく中で、例えばカラーの画像に関しては、フォトCDを一時期使っていたんですが、コダックさんがつくっていた規格ですけれども、これはもう今は

実際にはサポートが全くなくなってしまって、非常に再生が困難になっているというような形で、やはり技術の変化でせっかくつくった画像が使えなくなるということは、実際に発生しています。

またフィルムに戻ればいいという話もあるんですけども、やっぱりそこはコストがかかるので、デジタルで1回つくったものは、できるだけ有効に活用できるような流れというのは必要なかなと思っています。

もう一つはCD-ROMのほうですけども、こちらはやっぱりOSとかが再生環境がどんどん変わっていくことによって、非常に再生が困難になっていくということが次々に発生しています。これはほんとうにどう残すのかというのは難しい問題で、こういったパッケージ系のメディア自体が一時的なものなのかもしれませんが、やはり日本の文化を残すという意味では、非常に大きな問題だなと考えているところです。これはちょっと、先日報告書なども出しましたけれども、まだまだ調査研究が必要だと考えているところです。

【杉本座長】 すいません、基本的には90年代から開始されていると思うんですけども、要は十数年前に最初使い出した技術が、今でもそのまま使えているということでしょうか。

【大場委員】 そのままではないですかね。画像の変換などは途中でしていますので。ただ、例えば近代デジタルライブラリーなどを見ていただければ、初期にデジタル化したものは白黒の二値なんです。グレースケールというか、あまり見ばえがよくないと時々言われていますけれども、そういった利用者側のニーズというのが変わってきていて、求められている解像度とか、画像の品質が変わってきているので、それに対してはどう対応していくのか、じゃ、もう一回撮り直すのといった問題は実際発生しています。

【杉本座長】 どうもありがとうございます。

【山崎委員】 追加でいいですか。ちょうどパッケージ系の電子出版物と調査は、私がちょうど国会図書館時代に一度担当して書いたもので、つけ加えますけれども、ちょうどあの調査をやったときに、国会図書館の資料のパッケージ系のフロッピーだとかCD-ROMなどを、抽出調査をしたわけです。年代ごとに10年間ぐらいの抽出をしました。それは点数はそんなに多くなかったんですけども、数十点レベルで。

ちょうど1点1点、要するにパソコンに入れてみて実際に見えるかどうか、簡単に言えばそういう調査です。で、どこまで見られる。その起動で終わってしまうのか、あるいは起動はされるけれど中身は見られないとか、そういうさまざまなレベルの調査はしたんで

すけれども、実際見てみるとPDFと呼ばれている、ある程度安定的と思われるそのフォーマットでさえ見られなかったケースもありますし、それからOSの変換点で非常に増えていくんです。徐々にももちろん過去にさかのぼるに従って、当然中に見られないものが増えてきますが、10年さかのぼると8割、7割というレベルになっている。

ですから、パッケージングというのは非常に危険だなということと、それからOSの切りかえというのは既に5年、6年でありますので、ウィンドウズ95のときには一番多くなったんです。見られる、見られないの境目ですね。それからもう一つNECの98などという、あの当時はまだウィンドウズがない世界のところのは、全く見られないというのがあったんです。そういうものはかなり実際のところでありましたので、おそらく図書館でも博物館でも、既に所蔵されているものの過半数は見られないんじゃないでしょうか。パッケージングのものですね。ディスクに入っているものはもちろん問題ない。それでディスクに入れることを検討し始めたというのはそこにある。

【杉本座長】 どうもありがとうございます。ほか、あと5分ほどで終わりの時間ですけども。はい。

【盛田委員】 民間企業でデジタルアーカイブのお話を、中小の自治体さんから受けることが結構あるんですけど、そのときに1つはやっぱりデータをつくる、もしくはそれを配信するといった場合に、長期利用などの保存以外、保証というのを求められたりします。要するに、未来永劫データは見られるのかとか、未来永劫このサービスをしてくれるのかということに対しては、なかなか答えられない面があるんですけど、かといってそういったものをすべて、例えば国とか自治体がやるのかということのも、なかなか難しい状況もあるんでしょから、例えばサービスができなくなった場合は、そのデータを何らかの形で保証するとか、いわゆるE-JOURNALのバックナンバーみたいな考え方で、そういう受け皿というのも1つあってもいいのかなというのがありますし、かといって民間そのものがこうやっているいろいろやるということ自体も、非常に市場を活性化するという意味ではいいと思いますので、官と民がうまく連携できるような形ができると、非常にいいのかなと思います。

【杉本座長】 そういう意味で、ちょっとメーリングリストに出ていましたけれども、うまくダークアーカイブが使えるのかとか、そういうことになるのかもしれないですね。いかがでしょうか。どうぞ。

【田中委員】 実験はこれからやってみなければわからないんですが、この都道府県立

図書館の中間集約モデルというのは、非常に興味深い実験だと思うんですけど、ちょっとあれなのは、例えば国会図書館の側とかから見れば、都道府県単位で集約していただけると非常に大きなメリットがわかるんですけど、逆に都道府県の側で集約することに対してインセンティブが働くんだらうかというところで、これが広がるかどうかというところでは、もう少し倫理的なところで、これを実証することについては非常に有意義だと思うんですが、先々の展開というところで、やっぱり都道府県立側にも集約をすることがどういう意味があるのかということ、それをまたどうやってサポートできるのかというところは、ちょっと念頭に置いておいたほうがいいのかと思います。

【杉本座長】 どうもありがとうございます。調査のところで反映していただければと思います。

【常世田委員】 今のその地方のインセンティブの話にも関連してくるんですけども、このアーカイブを考えるのは当然のことながら、中央にあるような文化財的なものだけに注目するのではなくて、地方のいわばありきたりなものまで含めてアーカイブしていくというのが、前提としてあると思うんです。そうすると、今お話にあったような、インセンティブはだれが感じるのかということにもつながっていくわけです。

ここにいらっしゃる方たちは、人的な採用制度も確立されたような施設の方が多いので、地方の図書館とか美術館という状況になると、専門職の採用もまともに行われていないという状況がありますので、そうすると、今お話いただいたようなインセンティブを感じる人間自体が存在していない。

それからアメリカのアメリカン・メモリーなんかを見ても、非常にありきたりなものもどんどん収集していく。今はわからなかったけれども、10年、20年、50年後には非常に重要な価値を持つようなものも、当然その中に含まれていく。そういうものをやっぱり見きわめていくような専門職が地方にいないと、地方でだれがこの事業を関連的に行っていくかという問題になってくると思うんです。

このガイドラインの33ページの第3章の2の人的資源と教育というところ、今の段階では非常に文章的にもまだ十分ではないと思うんですけども、やっぱり地方で専門職をどうやって確保していくかというようなことを、ある程度提案していく必要があるんじゃないかなと思います。テクニカルなものについては、何をやらなきゃいけないかというのは比較的わかりやすい、解明されていくわけですけども、それをだれが担うのかということについては、やはりある程度の提言が必要なんじゃないかなと思います。

【杉本座長】 どうもありがとうございました。そうすると、大体時間が終わりにになりました。

ちょっと私、最後に一言つけ足したいのは、最近某博物館に行くことがありまして、そこで要はデジタルじゃないんですけれども、例えば2インチのビデオテープに入っているコンテンツをどうしたらいいんだろうとか、そういう電子的なものというのは、デジタルに限らずアナログでもいっぱい残っていたりして、その保存というのもきっと大きな課題になっているんだろうと思います。それも権利関係とかどうしたらいいのかという話もありますので、そうしたことも含めて考えていければなと思っております。

というところで、時間になりましたので、本日はどうもありがとうございました。

最後、事務局から。

【松田情報流通振興課統括補佐】 次回会合につきましては、追って事務局よりご連絡させていただきます。

また、引き続きメーリングリストでの意見交換をぜひさせていただければと思いますので、よろしくどうぞお願いいたします。

【杉本座長】 本日はどうもお忙しい中お集まりいただきまして、ありがとうございました。では、引き続きメーリングリスト等でもよろしくようお願いいたします。どうもありがとうございました。

— 了 —