



日本の文化を デジタル世界に伝える



京都大学人文科学研究所・共同研究班

「人文学研究資料にとってのWebの可能性を再探する」=編

永崎研宣=著



樹村房

はじめに

日本の文化は多様である。日本発祥のものもあれば、そうでないものもある。外来文化を自らの風土に合わせて取り込んで自家薬籠中のものとするのもまた日本の文化の特徴であると捉えるなら、外来文化もまた日本の文化の一部と考えてもよいだろう。日本の中だけでも多様である上に、そのようなものも含むと考えるなら、日本の文化が含むべきものの多様性はとても大きな広がりを持つことになる。

日本を育んできたこの多様な文化をなるべくそのままに未来世代にも継承していくことは、将来において日本が自らのアイデンティティを確立していく上で欠かせない要素の一つであるように思われる。それでは、文化に関わるものをこれまでどおり大切に保管し続けていけばよいのかと言えば、それだけでは立ちゆかなくなるかもしれない状況が様々な方面から押し寄せつつある。とりわけ、デジタル媒体を中心としたインフラ（インフラストラクチャー）の確立と一般の人々も含めた広範な普及、そしてそれに基づく人々の情報への接し方の変化は、文化に関わるものをデジタルインフラに適切に載せていくことの必要性を大幅に高めつつある。この点については、程度の多少はあれ、実感している人は多いだろう。

さて、表面的な動向だけを見ていると、横文字やアルファベットの略称が飛び交うような世界は、専門の情報技術者のみに任せておけばいいような事柄であるようにも思われる。しかし、何を、どのようにして、デジタルインフラに載せていけばよいのか、という点について、言い換えるなら、デジタル世界に何をどう伝えるのか、という課題に答えるのは、技術者だけでなく、文化に関わるもの、特に資料についての十分な知識を持った人でもある。デジタル技術が深化すればするほど、資料の個別の状況はより丁寧にデジタルインフラに載せることができるようになる。逆に言えば、載せることが可能だった個別の状況が載っていなかった場合、それはないものとされてしまう可能性もある。そのようにして、文化に関わる知識を持った人、つまり、文化の専門家が活躍し得る場は徐々に広がりつつある。

それでは、文化の専門家の側でもデジタル技術に関する深い知識を持たねばならないのか、と言えば、必ずしもそういうわけではない。深い知識を持ちたい人はともかく、そうではない人が、これからデジタルインフラに文化に関するものを載せていきたいと思っているときに、何をどれくらい知るべきなのか。あるいは、どれくらい知っておいた方がよいのか。このことに関する指針が日本語圏では必ずしも十分に共有されていないように思われる。そこで本書は、文化の専門家が、文化に関する資料をデジタルインフラに載せようとする時に知っておくべきこと、知っておいた方がよいこと、についての基本的な考え方をまとめた上で、具体的な近年の情報技術に即してそれを少し深めて考えてみることを目指したい。

さて、このように言ってしまうと、デジタル技術に関する知識の押し売りが行われているようにみえるかもしれない。本であれば、紙媒体であれば、このような裏方に関する知識をいちいち気にする必要はないのに、と思ってしまう人も多いだろう。たしかに、紙媒体の場合には、すでに大きな流通の仕組みが構築されており、それを通じて何かを伝えたいと思う人には、それぞれの状況に応じて様々な手段が用意されている。とりあえず原稿用紙に殴り書きして編集者に渡せばきれいな本になって国立国会図書館に永続的に保管される上に各地の書店や図書館を通じて読者のところに届き、刊行にかかった各種費用も回収できる。これをエコシステムと呼ぶこともあるようだが、文化を伝えるこの大きな仕組みができあがるのにどれほどの時間と労力が費やされたのか、出版流通史をひもとくだけでその膨大さに圧倒されてしまう。

しかし、デジタル世界のそれは、エコシステムと言えるような仕組みには至っておらず、様々な試みが世界中で進められているところであり、デジタル世界に文化を伝えようとするなら、その未完成なエコシステムに種をまくという姿勢が、多かれ少なかれ、どうしても必要になる。そして、時々様子を見て、水や肥料をやらないと、いつのまにか枯れてしまうということにもなりかねない。

ではデジタル世界に文化を伝えるのは時期尚早なのか、と言えば、それは各人の立場によるだろう。紙媒体のエコシステムも弱まりつつあるとは言え依然として根強いファンを多く抱え、仕組みとしてもまだ十分に機能していること

から、紙媒体にのみ注力し続け、そのよく練られた役割分担を十二分に活用するというあり方も当面は有効であると思われる。そのような紙媒体のエコシステムを構築してきた先人達と、同じか、あるいはそれ以上の努力が、今、まさに、デジタル世界での文化情報流通、あるいは知識流通のエコシステムにおいて必要とされている、というのが現在の状況なのである。

未来に向けて文化をよりよく伝えることを目指すなら、これから一層インフラとして定着し、広く情報を流通させるための主要な場になるであろうデジタル世界を、文化を伝える媒体としてふさわしいものにしていくという努力を誰かが担わなければならない。デジタル技術の専門家のみならず、そこには、伝えられる文化についてよく知る人々の力も当然のことながら不可欠であり、そして、紙媒体のエコシステムに携わってきた様々な関係者の知識と経験もまた、その将来像を描く上で重要になるだろう。

デジタル世界に日本の文化を伝えることは、それを通じて、世界中の人々に、そして未来の人々に我々の文化を伝えるということでもある。本書がそのことに多少なりともお役に立つことがあれば幸いである。

なお、本書が主に扱う対象は、これまで紙媒体で共有されてきた資料であり、音声・動画・3D等についてはあまり触れていない。そして、扱っている資料のほとんどは著作権保護期間が満了したものである。前者の理由については「あとがき」を参照されたい。後者については、それが現在は一番自由に幅広く使えるものであり、それ以外の資料のデジタル化と活用の際にも様々なヒントを提供できると考えたからである。また、本書では、同じような話が何度も登場することがある。これは、各章のテーマについてそれぞれの章の文脈において説明することを心がけたためであり、通読される読者には冗長なものと感じられるかもしれないが、ご容赦いただけるとありがたい。

日本の文化をデジタル世界に伝える

目次

はじめに iii

第1章 「デジタル世界に伝える」とは —————1

- 1.1 「デジタル世界に伝える」ためのおおまかな流れ1
- 1.2 デジタル世界に持って行けるものと持って行けないもの3
- 1.3 デジタル世界でのサステナビリティ5
- 1.4 デジタル世界に関する情報収集・共有8
- 1.5 まとめ9

第2章 デジタル世界への入り口 —————10

- 2.1 2値化・符号化の考え方10
- 2.2 どの程度まで詳細に記述するか12
- 2.3 事業としてのライフサイクル16

第3章 利便性を高めるには？ —————20

- 3.1 情報技術の発展への対応20
- 3.2 ユーザのニーズへの対応25
 - 3.2.1 技術の発展とユーザのニーズ 25
 - 3.2.2 ユーザの関心の深さの違いと利便性 28
 - 3.2.3 なるべく低コストで利便性を高めるために 36

第4章 デジタル世界に移行した後、なるべく長持ちさせるには —68

- 4.1 長持ちさせるための基本69
 - 4.1.1 必要な人材について 69
 - 4.1.2 必要なシステムと体制について 70
- 4.2 URLの維持73
- 4.3 システム移行による断絶について76

4.3.1	使い勝手の継承	76
4.3.2	データの継承	77
4.3.3	データを改良する場合	82
4.4	幅広い利用・活用に向けて	85
4.4.1	利用条件	86
4.4.2	コミュニティへの配慮	88
4.4.3	公開手法について	91
4.4.4	利用実績の提示	92
4.5	長期保存のための枠組み	93

第5章 可用性を高めるための国際的な決まり事：IIIF と TEI ———97

5.1	IIIF について	98
5.1.1	IIIF：問題とされたのは何だったのか？	98
5.1.2	IIIF：課題解決への取り組み	103
5.1.3	IIIF の現在	104
5.1.4	IIIF の広がり	114
5.1.5	IIIF コンテンツを活用するための仕組み	116
5.1.6	IIIF の導入に伴う公開の在り方の変化	117
5.1.7	IIIF の事例：SAT 大正蔵図像 DB	120
5.1.8	IIIF の事例：IIIF-BS（仏教研究のための IIIF マニフェスト集）	128
5.2	TEI について	137
5.2.1	デジタルテキストの特徴を活かすには？	137
5.2.2	TEI 登場のコンテキスト	139
5.2.3	TEI ガイドラインとは	141
5.2.4	TEI ガイドラインの活用事例	145
5.2.5	マークアップの深さをどう考えるか	157
5.2.6	テキストデータやツール・ノウハウを共有するには	158
5.2.7	どうやってマークアップするか	159
5.2.8	TEI に関するまとめ	162

第6章 実際の公開にあたって —————164

6.1 デジタル化について	164
6.1.1 デジタル化の対象の検討	164
6.1.2 デジタル化における個々の要素	167
6.2 デジタル化資料の Web 公開に際して	187
6.2.1 A1 書誌データ・目録データの公開	187
6.2.2 A2+A3 デジタル画像の公開	196
6.2.3 A4 画像アノテーション	202
6.2.4 A5 テキストデータの公開	203
6.2.5 A6 付帯情報の公開	208
6.3 まとめ	209

第7章 評価の問題 —————210

第8章 研究者の取り組みへの評価の問題 —————218

おわりに	223
さらに深めたい人・アップデートしたい人に	225
あとがき	231
索引	233

第 1 章

「デジタル世界に伝える」とは

1.1 「デジタル世界に伝える」ためのおおまかな流れ

「はじめに」で述べたように、デジタル世界に伝える、とは、現在のデジタルインフラを通じてつながる国内外のすべての人々、そして、それにつながるであろう未来のすべての人々に対してアクセスの可能性をひらくということである。そこで、まず、デジタルでないものであれば、デジタルデータに変換するというプロセスを経ることになる。そして、デジタルデータには適切なアクセス権を付与した上でデジタルインフラを通じて公開（あるいは非公開に）することになる。そして、いつでも誰でも共通の参照方法で一つのデジタル情報にアクセスできるようにしておく。

たとえば、国立国会図書館デジタルコレクションの『石見国絵図』は、10.11501/1286204というデジタルオブジェクト識別子（DOI）が割り当てられており、これを用いて DOI 財団のサイトにて <http://doi.org/10.11501/1286204> のようにしてアクセスすれば『石見国絵図』のデジタル画像を参照できる。そして、著作権保護期間満了と示されていることから、誰でもダウンロードして加工したり再配布したりできることになっている。

一つの資料をただデジタル世界に伝えるだけであれば、ここまでできていれば満足してしまうかもしれない。しかし、このままでは、アクセス可能性が開かれたとはいえ、この『石見国絵図』を閲覧したり利用したりしたいと思っている人のところに届くにはやや難しいものがある。たとえば、大きな書店や大きな図書館の書棚に配架されているだけでは見つけてもらうのが難しいということと似たような状況である。

しかも、紙の本であれば、背表紙の並んだ大型の書棚をざっと眺めてみるという所作を通じて見つけてもらえることもまれにあるが、現在のデジタルでは一覧性がそれほど高くないために、そのような仕方で見つけてもらえる可能性もあまり高くない。そこで、代わりに、デジタルの世界で情報発見のための主な手法になっている「検索」を活用するという選択肢が出てくる。

「検索」には、現在、いくつかの入り口がある。最大の入り口は Google や Bing、百度等の Web ページ全文検索エンジンである。こういった検索エンジンでは、クローラと呼ばれるソフトウェアを常時多数動かして、世界中の Web ページの情報を収集・蓄積し、いつでも検索できるようにしている。たとえばこういった検索エンジンに『石見国絵図』が登録されれば、利用者にとってはとても発見しやすいということになる¹。したがって、検索エンジンのクローラが自らのサイトの情報を収集・蓄積してくれるようにしておくこと、そして、その際に検索エンジンでヒットしやすいような用語の工夫等を行っておくことは一つの有力な選択肢だろう。

一方、自らのサイトで検索システムを提供する方法もある。それには、簡単なものから複雑なものまで、そして、無料のものから有料かつ高価なものまで、様々なソリューションが提供されている。使える予算や人手との兼ね合いで、どういうものを選ぶかは自ずと決まってくるだろう。検索の利便性としても色々な選択肢があるので、導入時によく検討されたいところである。

さらに、近年では、「統合検索」への期待感が高まりつつある。たとえば、各地に所蔵されている『石見国絵図』を探したいと思ったら、通常は、各地の図書館等の Web サイトにアクセスしてから検索することになる。しかし、現在のように多くの図書館が Web サイトを公開している状況では、その方法をとることは時間的制約からなかなか難しい。そこで、「統合検索」の出番となる。

「統合検索」の代表的なものに Europeana がある。この Web サイトでは、欧州の美術館・博物館・図書館・文書館等から5,800万件以上の資料の目録情報を収集し一括して検索できるようにしている。欧州各地に散らばっているも

1 : Google 検索結果の一例として、「デジタル岡山大百科」中の「石見国絵図」<http://di-gioka.libnet.pref.okayama.jp/detail-jp/id/kyo/M2014122011394969510>

のを一度に探すには大変便利な仕組みである。日本でも図書館資料にはNDLサーチ、大学図書館資料にはCiNii Books、博物館・美術館資料には文化遺産オンラインといった統合検索が広く提供されており、今後はジャパンサーチというプロジェクトによってさらに広範囲の統合検索が可能になると期待されているところである。昨今では、この種の統合検索に参加する、つまり、目録情報・書誌情報を提供して自らの資料群の発見性を高めることが、ユーザの利便性を高める重要な要素の一つとなっている。

以上が、デジタルインフラに文化資料を載せて、興味を持ってくれそうな人たちに使ってもらうための仕事のだいたいの流れである。このようにして書いてしまうと簡単のように思えなくもないが、実際には個々のプロセスで様々な工夫や努力が必要な場合があり、しかもそういった仕組みを支えるために背後で多大な労力が費やされていることも少なくない。そして、大抵の場合、労力を払った分だけ利用されやすくなり、対応に手間がかかる枠組みに対応すれば、やはりその分だけ利用されやすくなり、またその効果も大きなものになる場合が多い。しかし、この種の事柄に費やすことができる労力は有限である。担当者自身の努力であれ外注するための予算であれ、技術的に可能な最大限のことを実現するには十分ではないことが多い。場合によっては、人手も予算もそれほど多くない上にどんどん減らしていかなければならないことも少なくない。したがって、デジタル世界に文化資料を伝えるという行為は、そのような現実的な制約の中でできる範囲で、最大限の効果を目指して実行するということになる。

1.2 デジタル世界に持って行けるものと 持って行けないもの

忘れてはならないのは、デジタル世界に持って行けるものは、元の文化資料のすべてではなく、必ず何らかの取りこぼしが生じてしまう、という点である。労力をかければかけるほど、利便性を高めるだけでなく、取りこぼしを減らすこともできるが、それでも元の資料が持つ要素のすべてを持って行けるわけで

はない。あくまでも、デジタルに写し取れる範囲にとどまるのであり、さらに、様々な現実的な制約のために、「写し取ろうとしたこと」に限定されてしまうことになる。たとえば、羊皮紙資料のDNA情報²を残そうとするなら、ある程度のコストをかければその一部分を取り出して分析しデジタルにも記録することはできるが、資料中の別の箇所のDNA情報を確認したいと思ったとき、デジタル化されたデータでは為す術がない。あるいは、写本に使われた紙の材料を確認すべく高精度デジタル顕微鏡で撮影³したいと思った時、一つの写本資料の全箇所を撮影して顕微鏡レベルの超高精細なデジタル資料を作成し、写本のどの箇所でもデジタルで繊維の1本1本まで確認できるようになればデジタル世界に移すという意味では理想的だが、現段階でできることは一定数のサンプルをデジタル顕微鏡により撮影して提示しておくのが穏当であり、その他の箇所は諦めざるをえない。この場合は、少し遠い将来の技術の圧倒的な進歩により取りこぼしが減ることを期待するしかない。

デジタル翻刻（文字資料の文字をコンピュータに入力すること）について考えてみると、資料が汚れたり欠損したり形があいまいだったりして判断が分かれるような文字があったとしたら、それを文字として入力した時点で、その他の判断の可能性はテキストデータ上には残らなくなってしまう。このようなことであれば、国際的な記述ルール⁴がすでに用意されているので、それに対応するコストをかけられれば、翻刻者が気になった箇所に関しては、判断が分かれることについてもデジタル世界に移すことができる。とはいえ、翻刻者が気にせず入力してしまった箇所については、判断が分かれる余地があったかどうか、デジタルテキストからはわからなくなってしまう。あるいは、少し以前であれば、入力時の漢字をJIS第二水準までの範囲に限定するというルールの下、

2：羊皮紙のDNA情報の分析については2009年に論文が刊行されている。Timothy L. Stinson, Knowledge of the Flesh: Using DNA Analysis to Unlock Bibliographical Secrets of Medieval Parchment, The Papers of the Bibliographical Society of America, Vol. 103, No. 4 (DECEMBER 2009), pp. 435-453.

3：龍谷大学古典籍デジタルアーカイブ研究センターでは、フランス国立図書館が所蔵する敦煌出土写本の顕微鏡写真による紙繊維や混入物の分析を公開している。<http://www.afc.ryukoku.ac.jp/pelliot/indexjp.html>

4：ここでの記述ルールはTEI (Text Encoding Initiative) ガイドラインを指している。詳しくは第5章を参照されたい。

JIS 第二水準までに含まれない漢字は含まれるものに置き換えることで内容は同等に伝えたとするか、それが無理なら外字として注記するという形になっていたものをよく見かけた。この場合には元の資料がデジタルに移る時に文字の形状についての情報が多少欠損してしまった可能性を含むことになる。本質的には解決しようのない問題であるとはいえ、近年は Unicode で扱える漢字が圧倒的に充実してきたことにより、元の字形をなるべくそのままテキストデータとして表現できるようになってきており、問題はかなり解消されつつある。そして、他の諸問題と同様、コストをかければかけるほどこのギャップを縮めることは可能である。

デジタル世界に伝えるということは、あくまでもその時点の技術や規格でできる範囲で伝えるだけであって、つねに伝えるべき情報の幾分かは欠損してしまう。技術や規格が大幅に進歩したり、異なる観点からのデジタル化の要請が生じたりした場合には、改めて元の資料から伝え直す必要がある。したがって、未来世代に日本の文化を伝えようとするなら、デジタル世界に伝えるだけでなく、元の資料も可能なかぎり保存を続けることもまた重要であるということになる。ただし、これをもってデジタル化が単なるコスト増に過ぎないと直ちに判断するべきではない。デジタル化によって、現物資料に接触する機会を減らすことができるなら、汚損、劣化等の危険性を減らすこともでき、資料の保全性が高まるだけでなく、結果として資料の出納や修復にまつわる業務量を減らすこともできる。そして、それにもかかわらず、Web 公開できればむしろ閲覧される機会を大幅に増やすことができる。全体として考えた場合には、長い目で見ればコスト低減につながると見ることもできるだろう。

1.3 デジタル世界でのサステナビリティ

デジタル世界に伝えた後にも、それを継続的に利用してもらえるようにしなければならず、そこにもまたコストが要求される。公開に関わる機器やネットワーク、あるいは外部のサーバを間借りするといった形での費用が、デジタルインフラに載せた当初から発生することは少なくない。既存のものをうまく活用することでそれがすべて無料で済んだとしても、それを継続していくには、

また別の種類の困難さが生じてくる。維持するための運用費・運用者の確保である。これは、「公開」のように何か目新しい成果が出るのではなく、単に、誰かがアクセスしようとした時に、いつでもきちんとアクセスできるようにしておくだけである。とはいえ、放置しておけばずっと稼働し続けてくれるわけではなく、日常的なメンテナンスもさることながら、システムにセキュリティ問題が生じたり、ハードウェアが故障したりするなどした場合も考慮すると、やはり何らかの形でそれなりに対応できる人材が必要になる。しかし、このようなことに予算をかけるのは昨今の文化関連予算の縮減傾向からしても、なかなか容易ではないだろう。そして、それを数年間乗り切ったとしても、次に出てくるのはサーバやシステムの移行という大きな壁である。

サーバやシステムは、機器としての寿命があるだけでなく、ソフトウェアとしての寿命もある。ソフトウェアそのものがダメになってしまうわけではないのだが、ソフトウェアには、それを動かせるコンピュータ環境が必要である。たとえば、依拠しているオペレーションシステム（OS）のバージョンがあがってしまうと、動かなくなってしまうことがある。あるいは、近年の、特に資料等を公開するためのソフトウェアは、システムに潜むバグや不十分な仕様によってセキュリティホールが存在することが少なくない。通常は、それが発見された後、速やかにソフトウェアを開発する企業やコミュニティからセキュリティアップデートが公開され、各自のシステムにインストールできるようになる。しかし、ソフトウェアの設計が古くなってきたりすると、抜本的な設計変更が施され、便利で新しくなるものの、これまでのものからは直接アップデートできないものが開発され公開・販売されるようになる。その後しばらくすると、従来のソフトウェアのセキュリティ対策アップデートは行われなくなってしまう。そうすると、セキュリティホールが開いたままのソフトウェアを使い続けることはできないため、新規ソフトウェアに移行しなければならないことになる。ハードウェア自体の寿命やサポート期間の限界が5年程度であることが多く、これにソフトウェアの寿命も加わるため、結果として、多くの場合には5年程度で抜本的なシステム移行を迫られることになる。

システム移行のコストは様々である。ハードウェアからソフトウェアまですべてクラウドコンピューティングのような形で請け負ってくれる企業であれば、

契約内容次第では移行のコストはかからないということもあるかもしれない。しかし、なるべく自前で扱おうと思ったり、何らかの独自性を出すことを目指そうとする場合には、一定程度自前でソフトウェアやハードウェアを用意する必要がある。どこまで独自のものでまかなうかということは、そのプロジェクトの方針次第であり、ソフトウェアから自作することもできれば、企業のパッケージを購入してすべてお任せしてしまうこともできる。しかし、来たるシステム移行を考慮に入れるなら、この事柄は、後々に大きな影響を与える判断ということになる。では何かよい方法があるのかと言えば、残念ながら、少なくとも今のところ正解は存在しない。詳しくは後述するが、移行をスムーズにするための方策としてよく挙げられるのは、「ベンダーロックイン」等と呼ばれる特定企業の技術による囲い込みを避けて、なるべくオープンな標準規格・標準技術を採用しているシステムやソフトウェアやサービスを採用すること（あるいはそのようなサービスを提供している企業に依頼すること）である。特定企業のソフトウェア・システムでしか対応できないものになってしまった場合、他に選択肢がないために、移行時にかなりの費用を請求されても支払わざるを得なくなったり、あるいは、企業がそのソフトウェアの開発をやめてしまった場合、他社と一からまたやり直さなければならなくなることもある。いずれにしても、費用がかなりかかってしまうことになるため、「ベンダーロックイン」には十分に注意する必要がある。

また、たとえば、最近のMSワード（ソフトウェア）とそれを用いて作った文書ファイル（データ）の関係において、MSワードを新しいバージョンに入れ替えても古いワード文書を閲覧更新できるのと同様に、比較的汎用的な作成・アップデートが可能なソフトウェアに比べると、データはその内容に関する専門的個別的な知識を反映したものとして作成されるため、データの作成・作り直しはコストがかかることが多く、専門的であればあるほど、なるべく作り直しを避けるようにした方がよい。そうすると、データの形式をなるべく汎用的なものにしておいて、システムやソフトウェアの移行時にもデータはなるべくそのまま移行できるようにしておくという方法も有効になってくる。詳しくは後述するが、移行のコストに関しても、作成段階、あるいは改良する段階などでいくつかの対策が可能であり、文化に関する専門知識を持つ人々がそこ

に関わる余地も十分に用意されているのである。

1.4 デジタル世界に関する情報収集・共有

ここでもう一つ前提として確認しておきたいのは、この種の情報の共有が、なかなか容易ではないという点である。この点は、「デジタル世界に伝える」人だけでなく、デジタル世界に伝わったものを利用・活用する人にとってもなかなか容易ではない。

まず前者について考えてみると、このことについて適切な判断を行うためには、日進月歩のデジタル技術についてある程度の理解を有しつつ情報収集を継続的に行わなければならないが、その一方で、対象となる文化的なものについても一定程度の知識が必要となるからである。とりわけ、デジタル技術の発展の速度には驚くべきものがあり、書籍等の従来の情報収集手法ではついていくことがかなり難しく、Web上の情報源なども適宜参照していかねばならないという状況である。個別の技術動向についての情報収集が難しいだけでなく、そうした個々の動向が集まって織りなす全体の潮流も見えていかねばならず、これもまたわかりにくさを深めている。こういった情報を日々フォローしていく一方で、文化的なものにおいてそうした技術をどう応用・活用し得るかということについても配慮しておかねばならない。そのためには、扱っているものの特徴についてのある程度の知識が必要になることは言うまでもない。そのような原理的な難しさのなかでの判断が求められるのが、デジタル世界に文化を伝えるという営為なのである。

一方、後者に関しても、たとえばTwitterやFacebookなどのSNSを見ると、ほとんど毎日のように世界中のどこかから新たに文化的なものがデジタル世界に伝えられた（≡ Webで公開された）という情報が発信されている。利用者としての自分に深い関係があるものもあればそれほど深くないがいつかは役に立つかもしれない、本で言えば積ん読しておきたいような情報もあるだろう。とはいえ、その時々状況に応じて必要なものを常に適切に探し出すのはあまり容易なことではない。そういった情報を逐次集めてアップデートしてくれているWebサイトなどを把握しておくことが比較的現実的な対応だが、

それを十分に機能させるのはなかなか難しいことのように思われる。

しかも、規格や技術がアップデートされたり新規に普及したりして便利になりつつあるのだが、中には、その新たなルールを利用者の側でもある程度知っておいた方がよいということになる場合もある。本章の冒頭の例で言えば、DOIはその一例だろう。この枠組みにおいては、DOIのURLさえ提示しておけば、それでデジタルオブジェクトを長期にわたって識別できるとされていることから、DOIが付与されている論文情報や書誌情報のデジタル情報は、ブックマークするにしても、その情報について他の文書等で参照するにしても、アクセスしたURLよりもむしろDOIを記載しておくことが正しい振る舞いということになる。この事例に限らず、デジタル世界では、情報を適切に利用するために踏まえておくべき情報・常識・ルールなどが時々アップデートされてしまうことがあるため、利用者側でも対応を迫られる場合がある。そして、「デジタル世界に伝える」側からすれば、その新しくて素晴らしい何かを利用者にきちんと周知しなければならないということになる。

1.5 まとめ

ここまでで確認してきたことは、デジタル世界に文化を伝えていくことは常にコストと利便性のバランスの中で実現されていくものである、ということである。そうであれば、具体的には、何をどのようにすればどのように良くなり、どのようにコストがかかるのか、という点を把握しておくことが鍵になる。本書が目指すのは、これを読者の方々になるべく把握しやすい形で提示することで、デジタル世界に文化を伝えていくという営為を少しでもよりよいものにしていくことである。

第 2 章

デジタル世界への入り口

2.1 2 値化・符号化の考え方

デジタル世界とそうでない世界との間には、いくつかの乗り越えがたい壁がある。触れることのできる固有の現物資料であるかそうでないか、というのはすぐに思い浮かぶ相違点だが、これに関して言えば、たとえば、私たちはモナリザに触れることはできないし、そうでなくとも国宝級の文化財に触れることができる機会はきわめてまれであり、離れたところから視覚で体験することしかできないという場合も少なくない。また、研究でよく利用される古文書・古典籍等の場合には、影印本と呼ばれる、写真製版された本として刊行されることがあり、現物資料の代わりにこれを閲覧・利用することもある。デジタル世界においては現物資料に触れることができないというデメリットはあるものの、現物資料だからと言っていつも触れることができるというわけではなく、資料によってはまったく触れる機会を得られない場合もある。そのように捉えてみると、この壁は、デジタルかどうかという境界とは少しずれたところに以前から存在しており、乗り越えが困難であるかどうかはともかくとして、私たちはデジタル以前からその壁への対処の仕方を織り込みながら文化資料に接してきたと言えるだろう。

乗り越えがたい大きな壁の一つは、デジタルの世界では情報が2 値化されてしまう、という点である。ただし、ここで注意しておかねばならないのは、2 値化、すなわち、1 と 0 の信号に置き換えられる、と言ってしまうと、あたかも文化的なものが機械的に1 と 0 に置き換えられてしまうという印象を持ってしまいがちだが、実際には、その置き換えのルールは人間が決めている、とい

あとがき

本書は、京都大学人文科学研究所共同研究班「人文学研究資料にとっての Web の可能性を再探する（2013-2015年）」での議論とその後継続した検討内容の一部を反映したものである。共同研究班全体としては、人文学研究資料を Web で共有するための手法のみならず、それを活用して新たな知見を見いだすための技術や研究についての議論も多岐にわたって行われたが、本書で扱った内容は、その一部、主に前者となっている。後者については別の機会に公表することとしたい。また、音声・動画・3D 等の扱いが特に薄くなっているが、これは、それらを扱うための手法が十分にコモディティ化していると言えない状況であり、本書で扱ったとしても、それほど遠くない将来に大きな改訂が必要になる場合があり得ると判断したためである。

本書は共同研究班班長である永崎がすべての文章をとりまとめて書いたため、記述内容の誤りや偏りについてはすべて永崎に責任があるのでその点ご留意いただきたい。本書の執筆にあたっては、共同研究班メンバー、及び共同研究班の研究会やシンポジウムに登壇してくださった方々、さらにそれ以外にも多くの方々との議論が反映されている。とりわけ、東京大学の下田正弘先生及び吉見俊哉先生、国立情報学研究所の武田英明先生、高野明彦先生及び北本朝展先生、京都大学の原正一郎先生、近畿大学の田窪直規先生、東京国立博物館の田良島哲先生、キングスカレッジ・ロンドンの Harold Short 先生、ヴァージニア大学図書館の John Unsworth 先生、ヴィクトリア大学の Ray Siemens 先生、マーストリヒト大学の Susan Schreibman 先生、ニューカッスル大学の James Cummings 先生、ウィートン・カレッジの Kathryn Tomasek 先生、ハーバード大学の Peter Bol 先生及び Donald Sturgeon 先生、英国ナショナル・アーカイブズの Pip Willcox 先生、英国図書館の Kristian Jensen 先生におかれては、お名前を挙げて感謝したい。加えて、筆の遅い筆者の仕事を長くお待ちくださった樹村房の大塚栄一社長には深く感謝したい。

本書の刊行に至るまでに、以下の研究助成金による活動の一部が反映されていることを感謝とともに記しておく。JSPS 科研費 JP15H05725, JP26284068, JP16K02492, JP16H03422, JP18H03576, JP19H00516, JP19H00526, JP19H01241。

●共同研究班メンバーリスト（現所属，五十音順）

岩崎陽一（名古屋大学）

クリスティアン・ヴィッテルン（京都大学）

宇陀則彦（筑波大学）

大向一輝（東京大学）

北岡タマ子（お茶の水女子大学）

後藤真（国立歴史民俗博物館）

高田智和（国立国語研究所）

苔米地等流（一般財団法人人文情報学研究所）

永崎研宣（一般財団法人人文情報学研究所：班長）

松田訓典（国文学研究資料館）

守岡知彦（京都大学）

安岡孝一（京都大学：副班長）

山田太造（東京大学）

[編者]

京都大学人文科学研究所・共同研究班

「人文学研究資料にとっての Web の可能性を再探する」

(2013～2015年)

[著者]

永崎研宣（ながさき・きよのり）

一般財団法人人文情報学研究所主席研究員

1971年生。筑波大学大学院博士課程哲学・思想研究科単位取得退学。

博士（関西大学・文化交渉学）。

東京外国語大学アジア・アフリカ言語文化研究所 COE 研究員，山口県立大学国際文化学部助教授等を経て一般財団法人人文情報学研究所の設立に参画。

これまで，東京大学にて人文社会系研究科人文情報学拠点非常勤講師及び情報学環特任准教授，国立国会図書館研究員（委嘱），国際日本文化研究センター客員准教授，東京文化財研究所客員研究員，沖縄県立芸術大学附属研究所共同研究員，国立国語研究所共同研究員等，各地の大学研究機関で文化資料のデジタル化と応用についての研究支援活動を行ってきた。

学会関連活動としては，情報処理学会論文誌編集委員，日本印度学仏教学会常務委員情報担当，日本デジタル・ヒューマニティーズ学会議長，TEI Consortium 理事等。

<https://www.dhii.jp/nagasaki/>

日本の文化をデジタル世界に伝える

2019年9月10日 初版第1刷発行

検印廃止

編者◎ 京都大学人文科学研究所・共同
研究班「人文学研究資料にとっ
ての Web の可能性を再探する」

発行者 大塚 栄一

発行所 株式会社 樹村房

〒112-0002

東京都文京区小石川 5-11-7

電話 03-3868-7321

FAX 03-6801-5202

振替 00190-3-93169

<http://www.jusonbo.co.jp/>

印刷／亜細亜印刷株式会社

製本／有限会社愛千製本所

ISBN978-4-88367-327-8 乱丁・落丁本は小社にてお取り替えいたします。