

特集

分類 新時代

Digital Cultural Heritage における分類と 新しい情報知識体系の可能性

研谷 紀夫

はじめに

さまざまな文化資源をデジタル化したコンテンツについて、分類に焦点を絞った議論が行われることは少ない。そもそも文化資源に関するデジタルコンテンツを対象とした論文自体がそれほど多くはなく、その大部分はデジタル化に関する方法論や、メタデータに特化した議論である。そのため分類に焦点を絞った議論は部分的に指摘されるに留まっているのが現状である。

しかし、資料検索において、何らかの分類を活用することは、その利便性を高める重要な鍵となる。そのため、文化資源をデジタル化したコンテンツに関する分類の議論をより深める必要があるだろう。本稿ではデジタルミュージアム、アーカイブ、ライブラリーなどにおける分類の現状を踏まえた上で、三者を包括するにあるコンテンツを Digital Cultural Heritage として捉え、芸術系資料などを対象とした分類と新しい情報知識体系について考察する。

2 デジタルアーカイブ、デジタルミュージアム、 デジタルライブラリーと資料の構造化

近年、文化資源をデジタル化したデジタルコン

テンツに関する事例が多く発表されている。日本では1990年代より絵画・工芸品や各種の文書などのいわゆる歴史学や博物館学的に定義するところの、一次資料をデジタル化して公開するコンテンツが制作されている。これらはデジタルアーカイブという名称で呼ばれることが多いが、文書や記録資料などを資料群全体として捉えていく、本来のアーカイブとはやや異なった対象や方法論でまとめられたコンテンツである。

例えば、1998（平成10）年度通産省（当時）が行った「先導的アーカイブ映像制作支援事業」で制作された代表的なコンテンツは『21世紀に伝えたい日本の美術「狩野派」400年の歴史』『加賀前田家の工芸遺産～百工比照』『電子セット「幻の江戸城」』『日光東照宮～美の構造』『「国宝源氏物語絵巻」高精細デジタルアーカイブの試作』といったものであった¹⁾。

これらは美術作品を対象として、主に作品の審美性を多くの人に鑑賞してもらうエキシビジョン型のコンテンツであり、アーカイブというよりもミュージアムという語を使用したほうが適切なコンテンツであったと考えられる。しかし、当時の日本においては、アーカイブという語が一般には浸透していなかったため、デジタルという語とともに組み合わせて使用することで、新しいコンテンツを提示する言葉として用いられた側面があった。そのため、本来の意味をあまり顧みることがないままアーカイブという語が使用されてきたのが実情であろう。そこで、ここではアーカイブという語をより本来的な意味で捉え直してみた

とぎや のりお：東京大学大学院情報学環社会情報研究資料センター

キーワード：Digital Cultural Heritage

い。

アーカイブとは、何らかの記録に関する資料一点一点を個別にとっていくのではなく、資料全体でその構造と内容を捉えて情報化を行い、資料の読み解きから、人や組織の活動やさまざまな出来事などのいきさつを全体の文脈で捉えることができるような資料群およびその収蔵施設と定義する。この場合対象とする資料も、記録資料などが中心となるものの、上記のような定義に該当する資料であれば、美術品、図書、映像、その他のアート作品も含まれると考えられる。したがってデジタルアーカイブも、上記のような目的で資料の構造化や情報化を行って構成されるデジタルコンテンツであると考えられる。

デジタルライブラリーについては、さまざまな定義があるが、一つ一つが独立した刊本や雑誌・論文をデジタル化し、主に刊本の内容に即して分類・構成を行ったデジタルコンテンツと定義する。デジタルアーカイブとライブラリーの違いは、アーカイブは実際の人間活動の経緯や歴史を把握するために、より包括的な資料群の構造化をするのに対して、ライブラリーは一点一点の個別の内容を重視し、NDC（日本十進分類法）に代表されるような、内容を反映した包括的な分類体系に適応できる資料群ということができよう。

ミュージアムは、自然科学、考古学、文化・芸術、歴史の各系統において、ある一定の価値体系や視点、動機に基づいて、対象となる事物・事象を切り取り、その切り取る枠組み自体がその社会の文化や学問体系を反映し、それに基づいて学術・研究などが成立するような歴史的、文化的背景をもっている。これらミュージアムにおいては、視覚や聴覚など感覚器を総合的に用いて資料を把握するような作品・資料が多いことも指摘することができる。これらを踏まえると、ミュージアムの資料をデジタル化したデジタルミュージアムは、上記のような資料の構造化を、デジタル技術を用いて新たな視点で実現するコンテンツとして捉えることができよう。

しかし、一方で新しいデジタル技術は、こうした従来のミュージアム(M)・ライブラリー(L)・アーカイブ(A)などの枠組みを越境させる可能

性を秘めている。近年「MLA連携」という言葉が様々な所で見聞されるようになったが、これらも各館における資料のデジタル化とネットワークへの対応などが進展していることが一因であろう。

デジタル化された資料はネットワーク上で横断検索をすることが可能であり、Dublin Coreなどのような汎用的なメタデータのエレメントセットを用いることによって、ミュージアム、ライブラリー、アーカイブを横断的に検索できる環境も実現している。

既存の機関の連携ではなく、新しいデジタルコンテンツを構築する場合に、前述したようなMLAの特性のすべてを捉えたコンテンツを構築する場合もある。デジタル技術の特性を活用すれば、情報化の仕方やその蓄積の仕方によって、ある観点からみるとアーカイブ的な側面、またある観点からみるとライブラリー的な構成法から各資料にアクセスしていくことが可能である。また、表示・閲覧技術などを重視しながら、多様な歴史観や世界観を構成していくミュージアム的な世界を見せることも可能である。さらに、これらから捉えられない新しい資料の見せ方や資料の構成法も考えることができよう。

著者はこのようにデジタルミュージアム、ライブラリー、アーカイブの特性を包摂し、さらに既存のMLAの特性を越えるような資料構成や世界観などを提示させるようなコンテンツをDigital Cultural Heritageと呼称している。この言葉の定義も使用する者によって異なる側面があるが、アーカイブ、ミュージアム、ライブラリーといった概念から少し離れ、文化資源をデジタル化して公開するコンテンツとして捉えている点で共通しているといえよう。

このDigital Cultural Heritageについて論じた最近の動向としては、Fiona CameronとSarah Kenderdineらが主にメディア論などの立場から書いた書籍²⁾や、Lindsay MacDonaldによる技術的かつ実践的な観点から述べた書籍などをあげることができる³⁾。本稿では、このDigital Cultural Heritageという概念を、3機関の要素を包摂しながらも、デジタル技術を用いることによって新しい可能性を含んだ言葉として捉え、これらのデジ

タルコンテンツに関する分類について考察を進めていきたい。

3 アーカイブ、ミュージアム、ライブラリーを包摂する視点

Digital Cultural Heritage は、既存のミュージアム、アーカイブ、ライブラリーの資料の構成法を内包することが可能な立場であることはすでに述べたとおりである。そのため本論では、その考察の前に既存の3機関のセグメントの中で構築される各コンテンツとその分類の状況について考察してみたい。

現状では、3機関が現物資料の保存・構成法をそのまま反映させて構築したデジタルコンテンツにおいては、既存の分類法などが反映されている。日本でいえば、東京国立博物館の「名品ギャラリー」を閲覧してみると、その分類には博物館展示の現場や現物資料の分類が反映されている⁴⁾。

一方で国立国会図書館の「近代デジタルライブラリー」などを見ると、十進法による図書分類が提供されており、図書館の分類法から資料をアクセスすることが可能である^{5), 6)}。また国立公文書館が構築する「デジタルアーカイブ」においては、国立公文書館における出所原則に基づいた分類法を適用し、デジタルコンテンツを分類・構成している^{7), 8)}。ここでは現物資料の構成・分類法を忠実に踏襲しているといえる。

これらの既存の機関の構成を反映したデジタルコンテンツを、統合的に検索できるポータルサイトなどにおける分類はどのようにになっているだろうか。代表的なサイトである「国立国会図書館デジタルアーカイブポータル PORTA」では前述した国立国会図書館近代デジタルライブラリーのデジタルコンテンツの他、国立国会図書館、また国立の各博物館に関する資料を検索することが可能である^{9), 10)}。これらのサイトのコンテンツを掲載できる「詳細検索」のコーナーが設置されているが、そこではNDLC（国立国会図書館分類表）やNDC（日本十進分類法）などの検索を行えるようになっている。しかし、文書館の出所原則に基づく分類や、各種博物館の区分・分類などから

検索できるわけではない。そのため、図書館が運営主体ということもあり、ここでは図書館における分類が基本になっている。よって、PORTAにおける分類では、図書を中心としてNDLCやNDCなどの分類法を適応しているコンテンツに限定して検索ができると考えられる。

同じように文化資源を中心としたデジタルコンテンツのポータルサイトのような位置づけにある「文化遺産オンライン」においては、文化庁の文化財区分に近い分類から資料を検索できるようになっている^{11), 12)}。ここでの分類は建造物、絵画、版画、彫刻、工芸、考古学、歴史、その他の美術、民俗、伝統芸能、工芸技術、文化財保存技術、史跡、名勝、天然記念物などの分類となっている。

また文化庁は各種の文化財指定を検索するデータベース「国指定文化財等指定データベース」¹³⁾を公開しているが、そこでは文化庁の指定分類を反映した分類で検索ができる。例えば、国宝と重要文化財は、「美術工芸品」と「建造物」に分かれており、「美術工芸品」は絵画、彫刻、工芸品、書跡・典籍、古文書、考古資料、歴史資料などに分かれている。その他に史跡名勝天然記念物、重要無形文化財、重要有形民俗文化財、重要無形民俗文化財、重要文化的景観などの区分がある¹⁴⁾。

前述した「文化遺産オンライン」と「国指定文化財等指定データベース」の分類は細部においては異なっているが、大枠の分類においては類似点がある。文化庁の国宝と重要文化財の指定においては、対象資料だけでみると、古文書、歴史資料はアーカイブ的な要素があり、書跡・典籍はライブラリーの貴重書庫の要素を、それ以外はミュージアムの要素を含んでいる。

このように見ると、現在の文化資源をデジタル化したデジタルミュージアム、ライブラリー、アーカイブやそれを包摂する Digital Cultural Heritage においては、基本的には現物資料における分類を応用させて適応させている。そのため、MLA 連携の観点から考えると、三つの分野の中のどれか一つが適応されることがあっても、三者の分類が併用されることや、融合して使用されるような試みはまだ多くは見られないのが現状である。

しかし、ミュージアム、ライブラリー、アーカイブの3機関のコンテンツを連携していくことが可能な Digital Cultural Heritage においては、それらを連携・複合させるような新たな分類体系が必要であると考えられる。現状では前述したようにさまざまなコンテンツを複合する可能性がありながら、図書分類のみ、あるいは博物館分類のみ、となるなど、現状の機関の分類体系をそのまま活用することに留まっている。そのため、より多様な資料を複合的に統合できるような多様な知識体系が必要であると考えられる。また、それらは単に分類ということではなく、さまざまな事物の系統や関連性を表現する「情報知識体系」となる必要がある。著者はこのような現状を踏まえた上で Digital Cultural Heritage で活用する情報知識体系の構築を行っている。次項ではそれらの事例を踏まえながら今後の Digital Cultural Heritage における「情報知識体系」について考察したい。



4 複数の分類を複合的に結びつけた事例

著者らのグループは2006年に、弥生式土器の発見者の1人であり、明治期に人類学・考古学の研究者として活躍した坪井正五郎に関する資料などを格納する「文化資源統合アーカイブ」^{15), 16)}を構築した。研究者の個人資料であり、「図書・雑誌」「草稿・原稿」「写真」「標本」「ノート」「絵画」など多様な資料が格納されていた。

これらをデジタル化して公開する場合、多様な資料に対してどのように包括的な分類体系を該当させるかという課題があった。例えば、研究者であるため多数の蔵書があるが、これらの検索をよりスムーズに行うためには、何らかの図書に関する分類体系を設けることが望ましい。また、全体は記録資料群としての側面があるため、資料を原秩序の状況に応じて構成する必要がある。さらに、人類学や民俗学に関する用語や、年代や地名に応じた分類なども必要である。そのため、ライブラリー、アーカイブ、ミュージアムなどで使用される各種の分類の上位オントロジを用いて複合的に包摂し、さらに新たな情報項目も取り入れた情報知識体系を構築した¹⁷⁾。

オントロジは元来ギリシャ語で“存在”を表す語を語源としており、存在論などの思想・哲学における用語として使用されてきた歴史を持つ。一方で20世紀に入り、知識工学などの研究分野においては、機械にさまざまな知識や情報を理解させるための知識の体系化手法としてもこのオントロジという用語が使用されるようになった。

しかし、知識工学の分野においても複数の種類のオントロジがあり、存在論などの哲学的な議論を踏まえて、世界に存在する多くの事物の体系化を目指す上位オントロジと、特定の産業や状況などに関する概念の統制を行うドメインオントロジなどさまざまなタイプのオントロジが存在する¹⁸⁾。

またさらに、1990年代以降のインターネットの普及により、WEBにおける情報を機械においても読み込みが可能なように、情報知識を体系化する手法としてウェブオントロジが登場している¹⁹⁾。これらを踏まえて著者はオントロジについては以下のような定義付けを行った。

- (1) 世界、あるいは特定の枠組みに存在するさまざまな概念・事物を対象として、その体系化を目指す。
- (2) 主述の関係や集合などの方法を用いて、さまざまな事物の概念・性質・意味を定義する。
- (3) 一定の規則に基づいたルール・言語で記述する。
- (4) 汎用性を目指し、再利用が可能な共有化された知識とする。

これらの定義をそれぞれのオントロジにあてはめて考えると、「上位オントロジ」と呼ばれる体系は、極めて広い世界を対象として(1)の事項を(2)から(4)の方法を用いて構築したオントロジである。またドメインオントロジにおいては特定の世界で(1)を目指して、(2)から(4)の手法を用いて構築されるオントロジである。またウェブオントロジなどは、(1)の主な対象をウェブ上のリソースと設定した上で、(2)~(4)の手法を用いて構築されるオントロジであると定義できる。ただしウェブオントロジの現状は、RDFやOWLなど(3)の記述言語の開発に重点が置かれており、それらを使用してどのように(1)を実現していくかが大きな課題となっている。

これらの中で、本稿で紹介する Digital Cultural Heritage において活用することを目的としたオントロジは、主に歴史資料に掲載された世界を対象としながら、(2)を用いて(1)を構築することを目標として構築されたものである。

この上位オントロジの適応に際しては N. Guarino の上位オントロジを基本として、主に「1. 場所空間, 2. 時間, 3. 具体物, 4. 抽象物, 5. 事, 6. 行為・作用, 7. 社会, 8. 現象」などの項目に分けて構成を行った²⁰⁾。この体系において、図書分類は「5. 事」の下に配置し、またアーカイブの観点から必要な、資料の出所などを表す空間情報は「1. 場所」の下に配置した。さらに時間情報に関するものは、「2. 時間」の下に配置し、ミュージアムやその他の機関とも関連するような資料ジャンルに関するものは「5. 事」の下に配置した。

図書館における分類の活用は、図1で表示されているように、日本十進分類法に基づく10の大分類項目をグラフで示し、さらにこれらの各円をクリックすると下層に移ることになる。またグラフの右上にオレンジのアイコンが表示されているが、これをクリックすると、該当するデジタル化された資料が表示される。同じように、資料の出所原則と関連する場所の表示に関しては、それぞれの場所に関する情報をまとめ、図2で示すように体系化を行っている。図3では、世界や日本における主要な地名を体系化してグラフで表現した。同様に図4は時代、さらに図5は資料のジャンルを示している。

このオントロジでは既存の分類だけではなく、資料に関係する人物や組織など幅の広い知識体系をオントロジの中に含め、それぞれを結び合わせた。例えば、人物と関連する図書分類、場所と組織、歴史的な出来事と場所、文化資源のジャンルと人など、さまざまな体系をより有機的に結びつけている。資料に関する人や組織に関する情報が結びつけられている例は図6などのようなグラフを構成し、閲覧者に提供される。

このように本プロジェクトでは多様な資料に資料分類を適応させるため、複数の分類体系を上位オントロジのような、より包括的な情報知識体系

の中に組み入れた。概して多様な資料が格納される傾向のある Digital Cultural Heritage においては、このように既存の分類からより複合的な情報知識体系を用いることが必要だと考えられる。



芸術系資料を対象とした 情報知識体系の可能性

前項で示したような Digital Cultural Heritage における複合的な情報知識体系を他の分野に活用していくにはどのようにすればよいだろうか。前項で示した例は、人類学者に関する資料など、アーカイブや歴史・人類学系のミュージアムに近い資料群であった。しかし、前述したように1990年代に「デジタルアーカイブ」という名称で構築された文化資源をデジタル化したコンテンツの主な対象は「デジタルミュージアム」と呼ばれるような、芸術・工芸作品を対象としたものであった。そのため、現在でも日本では「美術館」と呼ばれる文化・芸術系のミュージアムに収蔵される作品がデジタル化されて公開されることが多い。ここではこれら多くの芸術系資料に適応することのできる複合的な情報知識体系についてその具体的なモデルを構想する。

現在デジタルコンテンツとして公開されている芸術系の文化資源においては、画像を中心としたエキシビションを提供する側面が強い。しかし、それらの作品について十分な知識を得るためには、それらと関係する記録資料や、参考文献などと有機的に結びつけていく必要がある。それを実現するためには、MLA の分類を平面的につなげていくだけではなく、その資料と関わるさまざまな事物をさまざまな関係性で立体的に結びつけていくことが必要であろう。

類似するサービスとして国立情報学研究所(NII)の連想検索などをあげることができる²¹⁾。これはすでによく知られているように、さまざまな事物・用語の遠近を分析し、それらの用語と関連する資料を検索していくシステムである。このシステムは、MLAの各機関に掲載された資料を機関のジャンルを超えてつなげていく長所をもっており、現在さまざまなコンテンツを連携させて

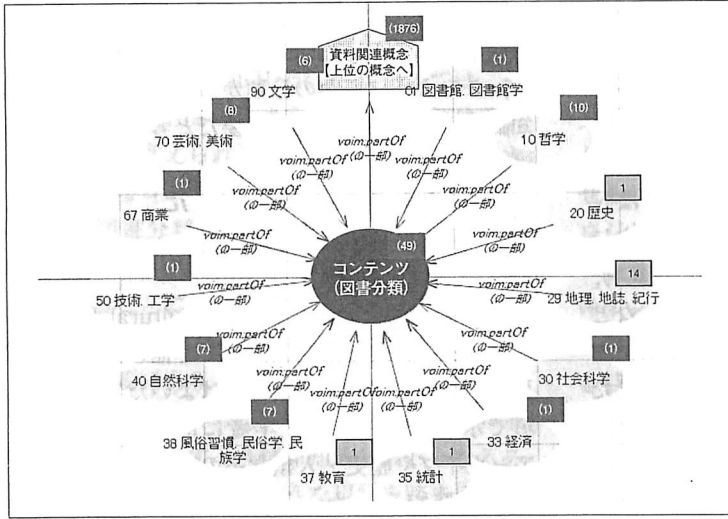


図1 図書分類の表示

■ はオレンジの色のアイコン
 これをクリックすると、
 該当するデジタル化した
 資料が表示される
 (図1, 2, 5, 6 共通)

図2 資料所在地に関する情報

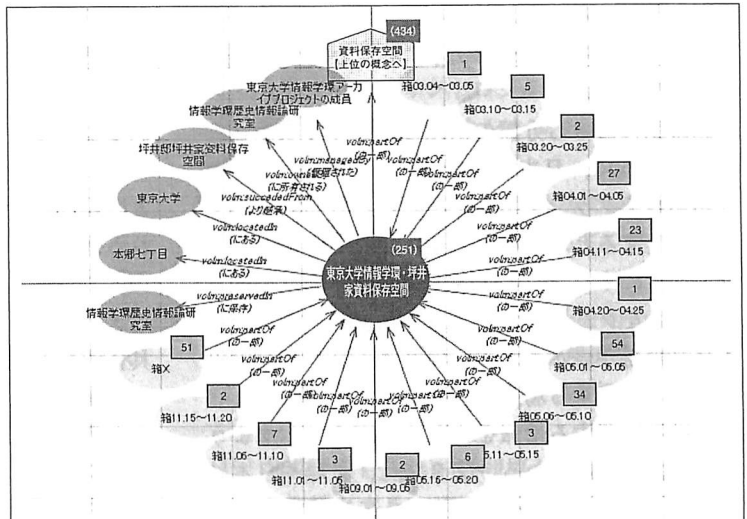


図3 地名に関する情報

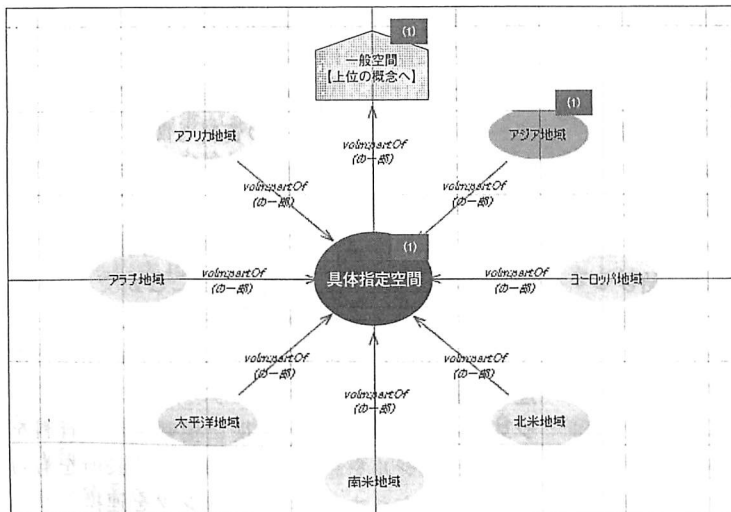


図4 時代に関する情報

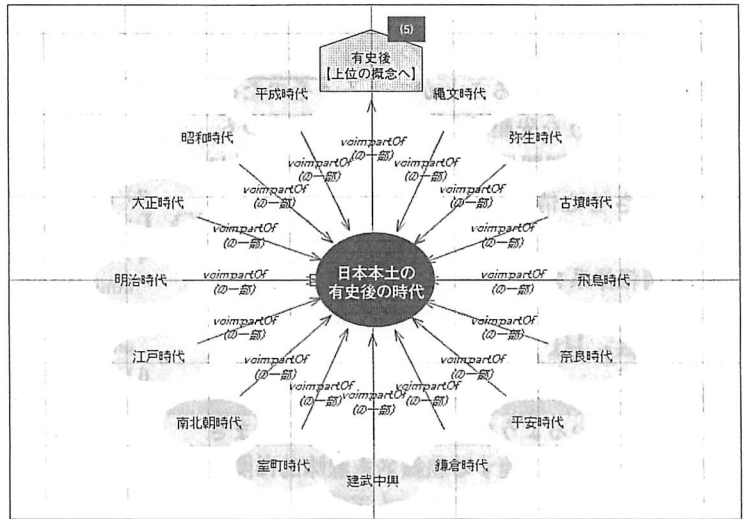


図5 資料ジャンルに関する情報

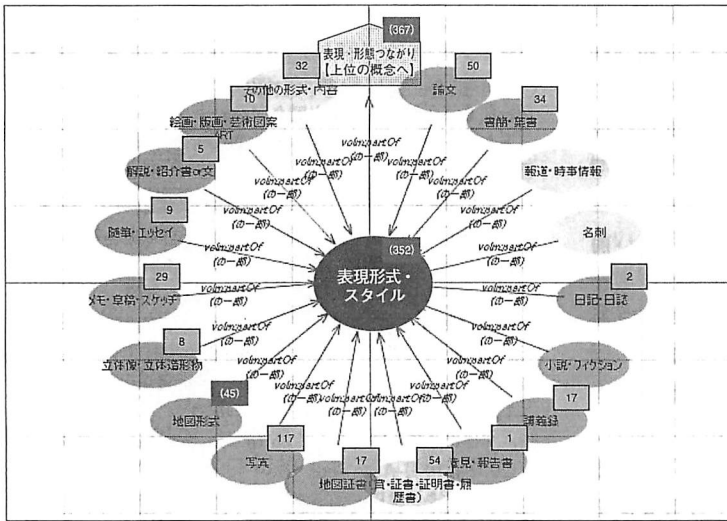
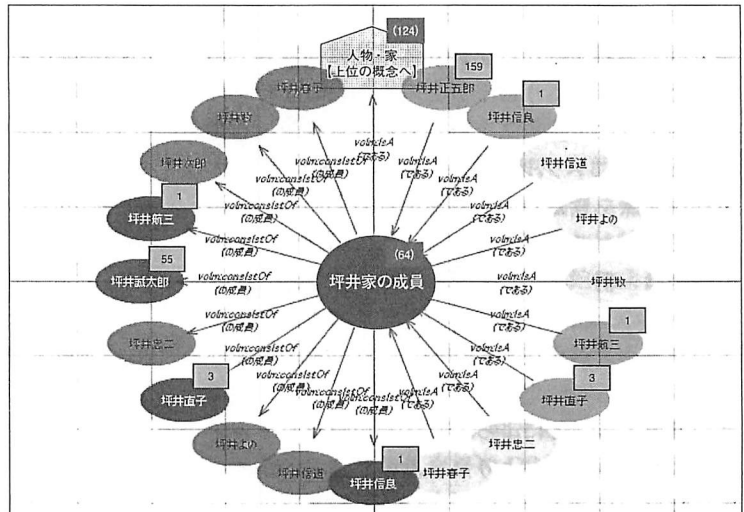


図6 資料と関連する人名と人物相互の関係についての情報



つなげている。しかし、より単語や概念、事物同士の相互関係が意味的に結びつけば、それらの内容をよりよく把握することができる。そのため、連想検索のような先駆的な成果を踏まえつつも、既存の分類や人間関係、地理情報などを意味的に相互連関させる情報知識体系を構築することによって、資料をさまざまな側面から捉えることができる情報体系を構築することが望ましい。

具体的には、芸術作品に関する文化庁の国指定文化財データベースや文化遺産オンラインなどの汎用的な分類に加えて、芸術家に関する人名事典に収録されているような、人名典拠情報やさまざまな流派や芸術と関連する組織名を意味的に関連づけることが望ましい。さらに、それらの中の、図書分類の中で関連する部分や、件名標目などを体系などを結びつけて、それぞれが相互に結びつくオントロジを構築することが必要である。このことによって、MLA 各側面の資料と、資料構成法をデジタルコンテンツの中で実現することができる。

例えば、明治時代の「黒田清輝」について、前述した国立国会図書館のデジタルアーカイブポータルで検索を行うと、各機関に収録された資料が、テーブルで996件程度表示される。これらについて、1件ずつ調べることも重要であるが、黒田清輝に関する人脈、関連組織、生誕地、時代、関連ジャンルからリンクされる各作品、関連標目・関連分類からリンクされる図書や資料などを結びつけることによって、黒田に関する作品、記録資料、参考文献などをより体系的に閲覧していくことができるようになると思われる。

このように、芸術系資料においても複合的な体系から資料にアクセスすることが可能となれば、多様な文脈を意識して資料を選択していくことが可能になると考えられる。これらの情報知識体系の構築は一朝一夕にはできないわけではないが、すでに芸術系の分類、件名標目、人名典拠などさまざまなリソースが存在するため、これらを公的な目的で活用していくことが重要であろう。このことが実現することによって文化資源をデジタル化して公開する意義がより高まるのではないかと考えられる。

<注>

- 1) 通商産業省機械情報産業局新映像産業室, 新映像産業推進センター, デジタルアーカイブ「先導的アーカイブ映像制作支援事業」報告, ニューメディア, 1999
- 2) Fiona Cameron Sarah Kenderdine, *Theorizing Digital Cultural Heritage A Critical Discourse*, mit press, 2007.
- 3) Lindsay MacDonald, *Digital Heritage: Applying Digital Imaging to Cultural Heritage*, Butterworth-Heinemann, 2006.
- 4) 東京国立博物館: <http://www.tnm.go.jp/jp/gallery/index.html> (2010年10月 Accessed)
- 5) 国立国会図書館: http://kindai.ndl.go.jp/BISmplSearch.php?tpl_bib_access=1 (2010年10月 Accessed)
- 6) 国立国会図書館, 近代デジタルライブラリー, 国立国会図書館月報, 587号, 2010, pp.14-18
- 7) <http://www.digital.archives.go.jp/> (2010年10月 Accessed)
- 8) 牟田昌平, 国立公文書館のデジタルアーカイブ: 過去の記録から未来の記憶へ, 情報処理学会研究報告 人文科学とコンピュータ研究会報告, 8, 2008, pp.17-24
- 9) 柴田昌樹, クローズアップNDL (第16回) 国立国会図書館デジタルアーカイブポータル (PORTA) について, 図書館雑誌, 103巻10号, 2009, p.708
- 10) 国立国会図書館, 国立国会図書館デジタルアーカイブポータル PORTA のご紹介, 国立国会図書館月報, 574号, 2009, pp.26-32
- 11) 文化遺産オンライン: <http://bunka.nii.ac.jp/SearchDisplay.do?linkType=keyword> (2010年10月 Accessed)
- 12) 丸川雄三, 文化財情報発信の実際—文化遺産オンラインの取り組みについて, 画像ラボ, 19巻4号, 2008, pp.26-29
- 13) 国指定文化財等指定データベース: <http://www.bunka.go.jp/bsys/> (2010年10月 Accessed)
- 14) 文化庁文化財保護部伝統文化課, 国指定文化財件数について, 月刊文化財, 356号, 1993, p.50-53
- 15) 文化資源統合アーカイブ: <http://cr-arch.chi3.u-tokyo.ac.jp/> (2010年10月 Accessed)
- 16) 研谷紀夫・津田光弘・倉持基・大島十二愛・山下大輔・松田好史・添野勉・馬場章, オントロジとコミュニティを用いた統合型デジタルアーカイブの構築の場合, 情報処理学会シンポジウム じんもんこん 2007 論文集, 2006, pp.57-62
- 17) 研谷紀夫・馬場章, 文化資源オントロジの構築とその活用, 情報知識学会誌, 17巻2号, 2007, pp.129-134
- 18) 溝口理一郎, 「オントロジー 工学 知の科学」, オーム社, 2005, pp.3-10
- 19) Tim Berners-Lee, "Semantic Web", *Scientific American*, 284 (5), 2001, pp.34-44
- 20) Nicolas Guarino, "Some organizing principles for a unified top-level ontology", Working Notes of AAAI Spring Symposium on Ontological Engineering, 1997
- 21) 丸川雄三, 阿辺川 武, 横断的連想検索サービス「想-IMAGINE」データベース連携が拓く新たな可能性, 情報管理, 53巻4号, 2010, pp.198-204

(2010.10.19 受理)