

- (27) "Author Resolver". RefWorks-COS.  
<http://www.refworks-cos.com/authorresolver/>, (accessed 2011-01-14).
- (28) "PubMed". National Center for Biotechnology Information.  
<http://www.ncbi.nlm.nih.gov/pubmed>, (accessed 2011-01-14).
- (29) "Digital Author Identifier (DAI)". SURF Foundation.  
<http://www.surfoundation.nl/en/themas/openonderzoek/infrastructuur/Pages/digitalauthoridentifiervai.aspx>, (accessed 2011-01-14).
- (30) "Author Identifiers". ArXiv.org.  
[http://arxiv.org/help/author\\_identifiers](http://arxiv.org/help/author_identifiers), (accessed 2011-01-14).
- (31) "Names Project". Mimas.  
<http://names.mimas.ac.uk/>, (accessed 2011-01-14).
- (32) "研究者リゾルバー". 国立情報学研究所.  
<http://rns.nii.ac.jp/>, (参照 2011-01-14).
- (33) "科学研究費補助金データベース KAKEN". 国立情報学研究所.  
<http://kaken.nii.ac.jp/>, (参照 2011-01-14).

## CA1741

## 動向レビュー

## 人文学研究と電子アーカイブ

## 1. 電子アーカイブプロジェクト

## 1.1. 人文学研究と資料アクセス

14世紀英国の物語詩『農夫ピアズの夢』("Piers Plowman")には3つの稿と十指にあまる写本がある。各バージョンを合わせると60以上の基礎資料が存在し、手稿のページ数は1万にのぼる<sup>(1)</sup>。

手稿や異稿の研究は人文学に欠かせないが、原資料に直接アクセスしたり、各地に分散する異稿を調べて回ったりするのは容易ではない。そこでこうした資料を整理し、注釈や関連情報とともに提供する学術版、批判校訂版、あるいはファクシミリ版(たとえば『農夫ピアズの夢』コンコードダンス)や手稿B.15.17ケンブリッジ版)が重要な役割を果たすことになる。

しかし印刷物では、物理的、経済的な制約から盛り込める情報が限定され、検索や相互参照などの活用に限界がある。そこで学術資料をデジタル化する試みが重ねられ、さらにインターネットの発達とともに、『農夫ピアズの夢』電子アーカイブ<sup>(2)</sup>のようなデジタル化されたアーカイブに発展し、資料へのアクセス性が飛躍的に向上してきた(電子アーカイブはデジタル・アーカイブとも呼ばれる。これらの場合のアーカイブは、資料の集成だけでなく、一般に学術版としての研究成果も盛り込んだサイトの意味で使われる<sup>(3)</sup>)。

## 1.2. 資料デジタル化の役割

資料のデジタル化は、膨大な資料へのアクセス性を改善するばかりではない。マッギャン(Jerome McGan)は、紙ベースのテキストを電子形態に変換すると原資料の見方が大きく変わることを指摘し、それは「自然現象研究に対する数学的アプローチが理論的視点のレベルを高度化すると同じように、電子ツールが批判的抽象度のレベルを引き上げる」からだと述べている<sup>(4)</sup>。

研究ツールとしてデジタルテキストを利用するためには、手稿や印刷物などの形の原資料から文字を転写し、さらにそのテキストがどのような構造になっているか(ページ構成、章節構成など)を何らかの方法で明示しなければならない。またアーカイブされた資料を検索利用し、共有するためには、メタデータを適切に付与することも重要である。

テキストのデジタル化に関しては、転写の方法、文字コードなど、問題となる点は多々あるが、ここ

ではテキストの構造を示すためのマーク付けとメタデータの表現が、主要な電子アーカイブでどのように行なわれてきたか、その現状と展望を概観する。

### 1.3. ペルセウス電子図書館での TEI 利用

「ペルセウス電子図書館」(Perseus Digital Library)<sup>(5)</sup>は、米国タフツ大学で進められてきた、電子アーカイブの中でも最も古い歴史を持つプロジェクトのひとつである。当初は CD-ROM としてギリシャ古典文学のアーカイブが出版されたが、1995 年には早くもウェブ版が提供され、扱う範囲もルネサンス、19 世紀米国などへと拡大してきている。

テキストのマーク付けには、TEI 規格<sup>(6)</sup>が採用された。TEI は、文学テキストから古文書や碑文まで、多様なテキストの構造とメタデータを記述するための国際標準で、章、段落、韻文の行といった文書構造だけでなく、文中に出現するキーワードを示すための要素も提供している。たとえば日付なら <date>、地名なら <placeName> で次のようにマーク付けできる。

```
<date value="-61" authname="-61">61</date> B.C.
<placeName key="tgn,7009327"
  authname="tgn,7009327">Marne</placeName>
```

こうして「61」という数字は「紀元前 61 年」を意味すること、「Marne」は地名であることが示され、さらに <placeName> の属性に地名シソーラス TGN のコードを加えることで「Marne」はドイツや米国ではなくフランス東部の地名であることが明示される。

ウェブ用には、TEI の XML を一定の規則で HTML に変換しており、通常のブラウザでそのまま閲覧できる。年代や地名としてマーク付けされた箇所はハイパーリンクで示され、コレクション内の資料で同じ地名や年代が出現する箇所を一覧表示するようになっている。

### 1.4. ロセッティ・アーカイブでの XML 記述

「ロセッティ・アーカイブ」(Rossetti Archive)<sup>(7)</sup>は、19 世紀イギリスの画家・詩人であるダンテ・ガブリエル・ロセッティ (Dante Gabriel Rossetti) の全作品をデジタル化して収録する、米国バージニア大学人文科学高度技術研究所 (IATH) のプロジェクトである。

ここでは絵画などの物理構造を中心にした記述や、文書構造と物理構造を混合した記述が必要になる。しかしプロジェクトが開始された 1990 年代前半においては、TEI は文書構造の記述が中心で、この目的に適さないと考えられた。そこでこのアーカ

イブでは、抽象的な「作品」を示す RAW (Rossetti Archive Work)、テキスト文書を表す RAD (同 Document)、絵画を表す RAP (同 Picture)、さらにジャンルを表す RAC (同 Commentary) という 4 つの文書型 (文書構造記述のスキーマ) を独自に定義した<sup>(8)</sup>。

現在これらは RAM (Rossetti Archive Master) という新しい文書型に統合され、各ファイルは <ram> 要素内に <ramheader> およびファイルの種類 (RAW、RAD、RAP) ごとの要素を持つようになっている<sup>(9)</sup>。たとえば有名な複合作品「召された乙女」("The Blessed Damozel") の RAW の場合、画面表示用説明を含むさまざまな情報が <ramheader> に収められ、並行して推奨テキストの RAD を示す <readingtext>、代表的な画像の RAP を示す <viewingimage>、「作品」を構成する個別の RAD、RAP を列挙する <wclist> を持つという形になる。

```
<ram>
  |-- <ramheader>
  |-- <readingtext>
  |-- <viewingimage>
  |-- <wclist>
```

その後 TEI は改定を重ね、デジタル化プロジェクトでの利用を念頭に置いた推奨記述法を公開するようになっており<sup>(10)</sup>、IATH の研究者たちも TEI を採用するほうが良かったかも知れないと述懐している<sup>(11)</sup>。マーク付けに標準仕様を用いるか固有の問題に最適化したものを開発するかは難しい選択だが、互いの情報を交換・共有するためには、特に目録あるいはメタデータに関してだけでも、共通の形で記述することが重要になるだろう。

## 2. メタデータの記述と応用

### 2.1. 電子アーカイブのメタデータ記述

TEI をはじめ各アーカイブの文書型は、一般にヘッダという形で記述対象のメタデータを保持できる。ただ TEI の <teiHeader> にせよロセッティにおける RAM の <ramheader> にせよ、多様な情報を柔軟に扱うことができる一方、複雑すぎて共有のためのメタデータとしては利用しにくい。実際、ペルセウス電子図書館も公開用 XML では <teiHeader> を省いてしまい、内部的に保持したメタデータをダブリン・コア (DC) として OAI-PMH で提供したり、MODS で表現したりする試みが行なわれている<sup>(12)(13)</sup>。

ロセッティ・アーカイブが参加する NINES (19 世紀電子学術研究のためのネットワーク化基盤) は、

各プロジェクトの資料検索を円滑に行なうために、RDF (Resource Description Framework) によるメタデータ記述を進めている。このメタデータは、DC、MARC21 の Code List for Relators、および NINES が提供するツール Collex で用いる語彙で構成される<sup>(14)</sup>。

ロセッティ・アーカイブの「召された乙女」の RDF メタデータを見ると、同作品の RAW ファイルから <ramheader> の主要情報、<readingtext>、<viewingimage> に対応するテキストおよび画像ファイル、さらに <wclist> に列挙される各アイテムが取り出され、作品を表す URI を主語にフラットに並ぶというシンプルな構造になっている。アーカイブごとに異なる構造を捨象して扱いを容易にするとともに、基本部分についてダブリン・コアを採用したことで、NINES に限らず広い範囲でのデータ交換、共有が可能になったといえる。

## 2.2. 電子図書館ポータルへのメタデータ

Europeana<sup>(15)</sup> は、欧州各国の図書館、博物館等が保有する文化遺産を横断的に検索できるようにするプロジェクトである。1,000 万を越す膨大な資源のメタデータを標準形式で集約し、さまざまな切り口での検索・表示を提供する「電子図書館ポータルサイト」として、個々の作品の研究結果を自ら提供する電子学術アーカイブとはまた異なる方法で、高度な資料アクセス性を実現している。

オランダのアムステルダム自由大学において、この Europeana のセマンティック検索エンジン<sup>(16)</sup> が試験運用されている。ここでは例えばキーワード“Bonn”に対して、それは地名なのか人名なのか「ボンの征服」というテーマを扱ったものかといった選択肢が示され、地名としてのボンを選べばその TGN コードに基づく URI を用いて検索が行なわれる。検索結果は、ボンを描いた作品、ボンで生まれた作家による作品など、ボンとどのような関わりがあるかによって細かく分類して表示される。

こうした検索が可能になるのは、作品のメタデータが RDF によって記述され、さらに作者についても birthDate、birthPlace といったプロパティでメタデータが用意されているからだ。これらのプロパティ値は URI で表現されており、検索を確実にするだけでなく、URI をたどって RDF データがつながる「リンクするデータ」(Linked Data) も目指されている<sup>(17)</sup>。

## 3. HTML と電子学術アーカイブ

### 3.1. マーク付けの変遷

電子学術アーカイブは、まず単純な HTML によ

る試験的サイトが公開され、第 2 段階で XML による記述に移行しているケースが多い。HTML では細かなマーク付けやそれを利用した高度な検索が難しいこと、XML であれば柔軟な情報表現が可能な上に、表示用ページも XSLT (元の XML 文書を別の XHTML などに変換する技術) などのツールで体系的に生成できるといった理由による<sup>(18)</sup>。しかし、多くのプロジェクトが独自の XML とそのマーク付けを前提にしたツールを開発した結果、アーカイブ間の相互運用性は低くなってしまった<sup>(19)</sup>。

一方で、HTML の属性を用いて文書内のデータを構造的に示す RDFa 規格<sup>(20)</sup> が W3C から 2008 年に勧告され、HTML でも詳細な情報記述が可能になってきた。ペルセウス電子図書館での例に挙げた TEI の地名マーク付けを RDFa による HTML の属性で表現するならば、次のような記述ができるだろう。

```
<span rel="tei:placeName">
  <a href="?tgn:7009327"
    property="rdfs:label">Marne</a>
</span>
```

RDFa でマーク付けした HTML からは、RDF の内容をプログラムで抽出できるので、メタデータファイルを別途用意する必要もない。Europeana の検索結果詳細ページには、RDFa を用いてダブリン・コアによるメタデータが埋め込まれ、ウェブブラウザ向けの情報と RDF 処理ツール向けの情報が一元的に提供されている。

### 3.2. コラボレーションと注釈

シリングスバーグ (Peter L. Shillingsburg) は、電子学術アーカイブにおけるコラボレーション、あるいは利用者による拡張の必要性を強調している<sup>(21)</sup>。そのひとつの要素が、アーカイブの資料に利用者がコメントや注釈を加える機能だ。ユーザ注釈が可能なアーカイブはあるが、多くの場合独自のツールを用いているため、サイトを横断したコラボレーションは実現できない。

アーカイブの資料ページを直接編集することなく、利用者が自由に注釈を加える手段としては、W3C のアノテア・プロジェクト (Annotea Project)<sup>(22)</sup> での試みが挙げられる。これはウェブ文書の特典箇所を URI と XPointer (XML 文書内の部分を示す手段) を用いて識別し、その部分に関する注釈、注釈者、日時などを RDF で表現して注釈サーバに保存していくというものだ。シンプルなオープン仕様なので、特定のツールに依存せず誰でもサービスを実装でき

る。W3C のサイトでテストサーバーが提供されており、クライアント用ツールとしても Firefox アドオンや Java によるブラウザなどが公開されている。

残念ながら、2011 年 1 月現在プロジェクトはあまり活発ではないが、それぞれのアーカイブが資料を分散公開しつつ利用者注釈などのコラボレーションを実現するための標準フォーマット候補として、注目しておきたい。

### 3.3. 韻文、縦書き、ルビ

HTML は単純な文書構造しか想定していないが、人文学研究においては韻文、戯曲などの記述も必要になる。日本国内では縦書き、ルビの表現も大きな問題だ。

次期の HTML5<sup>(23)</sup> では (XHTML と同様に) 名前空間を用いて他の語彙を組み入れることができる。基本マーク付け (ホスト言語) には HTML を使い、韻文には TEI の <tei:lg>, <tei:l> を使うといった組み合わせにより、利用しやすさと精緻な記述の両立が可能だ。

縦書き、ルビ表示は、これまでもできないわけではなかったが、実装が不揃いで利用しにくい状況だった。2011 年 1 月現在、電子書籍での日本語表示に関する仕様策定が進むのと並行して、CSS3 での縦書き正式採用に向けた準備が活発に行なわれている<sup>(24)</sup>。また HTML5 では仕様本体にルビ表示機能が取り込まれる予定で、主要ブラウザで縦書き、ルビ表示共に標準的に可能になる日も近い。

文学作品の優れた学術版が作られていても、使い勝手の悪さから、専門家ですらペーパーバックのような普及版に基づいて研究を進めてしまう場合があるとシリングスバーグは指摘する。HTML を基本にした電子アーカイブは、通常のブラウザや汎用ツールで検索やデータ抽出、再利用ができ、利用者にとっての利便性は高い。また専用の文書型やツールを開発する必要がなく、アーカイブ作成・公開のためのコスト、時間も大きく低下させることができるだろう。標準化と固有性の間で難しい選択を続けてきた電子アーカイブにとって、HTML によるマーク付けを改めて考え直してみる機会が到来している。

(kanzaki.com : 神崎正英<sup>かんざきまさひで</sup>)

(1) Fenton, Eileen Gifford et al. "Effective methods of producing machine-readable text from manuscript and print sources". Electronic Textual Editing. Burnard, Lou et al., eds. New York, Modern Language Association of America, 2006, p. 241-253.  
 (2) Piers Plowman Electronic Archive. <http://www3.iath.virginia.edu/seenet/piers/>, (accessed 2011-01-14).  
 (3) Price, Kenneth M. "Electronic scholarly editions". A Companion to Digital Literary Studies. Schreibman, Susan

et al., eds. Malden, MA, USA, Blackwell Publishing, 2007, p. 434-450.  
<http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405148641/9781405148641.xml>, (accessed 2011-01-14).  
 (4) McGann, Jerome. "Imagining what you don't know: The theoretical goals of the Rossetti Archive". Institute for Advanced Technology in the Humanities. 2010-07-14. <http://www2.iath.virginia.edu/jjm2f/old/chum.html>, (accessed 2011-01-14).  
 (5) Perseus Digital Library. <http://www.perseus.tufts.edu/hopper/>, (accessed 2011-01-14).  
 (6) Burnard, Lou et al., eds. "TEI P5: Guidelines for electronic text encoding and interchange". TEI Consortium. 2010-11-05. <http://www.tei-c.org/release/doc/tei-p5-doc/html/>, (accessed 2011-01-14).  
 (7) Rossetti Archive. <http://www.rossettiarchive.org/>, (accessed 2011-01-14).  
 (8) Pitti, Daniel et al. "After the Fall: Structured Data at IATH". Institute for Advanced Technology in the Humanities. <http://www.iath.virginia.edu/~jmu2m/ach98.html>, (accessed 2011-01-14).  
 (9) Rossetti Archive Master. <http://www.rossettiarchive.org/ram.xsd>, (accessed 2011-01-14).  
 (10) "Best Practices for TEI in Libraries". Text Encoding Initiative. <http://purl.oclc.org/NET/teiinlibraries>, (accessed 2011-01-14).  
 (11) Pitti, Daniel et al. "After the Fall: Structured Data at IATH". Institute for Advanced Technology in the Humanities. <http://www.iath.virginia.edu/~jmu2m/ach98.html>, (accessed 2011-01-14).  
 (12) Smith, David A. et al. "Integrating harvesting into digital library content". Proceedings of the Second ACM/IEEE-CS Joint Conference on Digital Libraries. Portland, OR, USA, 2002-07-14/18. New York, ACM Press, 2002, p. 183-184. <http://hdl.handle.net/10427/57022>, (accessed 2011-01-14).  
 (13) Babeu, Alison. "Building a "FRBR-Inspired" Catalog: The Perseus Digital Library Experience". Persues Digital Library. 2008-01-31. <http://www.perseus.tufts.edu/publications/PerseusFRBRExperiment.pdf>, (accessed 2011-01-14).  
 (14) "RDF samples". NINESWiki. [http://www.performantsoftware.com/nines\\_wiki/index.php/RDF\\_samples](http://www.performantsoftware.com/nines_wiki/index.php/RDF_samples), (accessed 2011-01-14).  
 (15) Europeana. <http://europeana.eu/portal/>, (accessed 2011-01-14).  
 (16) A research prototype of Europeana's semantic search engine. <http://eculture.cs.vu.nl/europeana/session/search>, (accessed 2011-01-14).  
 (17) Zeinstra, Maarten et al. "Open linked data and Europeana". Europeana. 2010-11-22. [https://version1.europeana.eu/c/document\\_library/get\\_file?uuid=374c381f-a48b-4cf0-bbde-172cf03672a2&groupId=10602](https://version1.europeana.eu/c/document_library/get_file?uuid=374c381f-a48b-4cf0-bbde-172cf03672a2&groupId=10602), (accessed 2011-01-14).  
 (18) Ellison, Lela et al. "Digital Humanities in early online archives". 2009-12-03. <http://hdl.handle.net/2152/6837>, (accessed 2011-01-14).  
 (19) 電子アーカイブの相互運用性が今後の課題であることについては、たとえばシリングスバーグが次の論文で取り上げている。  
 Shillingsburg, Peter. "How literary works exist: convenient scholarly editions". Digital Humanities Quarterly. 2009, 3(3). <http://digitalhumanities.org/dhq/vol/3/3/000054/000054.html>, (accessed 2011-01-14).  
 (20) Adida, Ben et al., eds. "RDFa in XHTML: syntax and processing". W3C. 2008-10-14. <http://www.w3.org/TR/rdfa-syntax/>, (accessed 2011-01-14).  
 (21) シリングスバーグ, ピーター. "書記行為を再現するための電子的インフラストラクチャー". ゲーテンベルクからグーグルへ: 文学テキストのデジタル化と編集文献学. 明星聖子ほか訳. 慶應義塾大学出版会, 2009, p. 105-170.  
 (22) "Annotea project". W3C. 2005-10-31. <http://www.w3.org/2001/Annotea/>, (accessed 2011-01-14).



- (23) Hickson, Ian. "HTML5: a vocabulary and associated APIs for HTML and XHTML W3C Working Draft". W3C. 2011-01-13.  
<http://www.w3.org/TR/html5/>, (accessed 2011-02-05).
- (24) Etemad, Erika J. et al. "CSS writing modes module level 3 W3C Working Draft". W3C. 2010-12-02.  
<http://www.w3.org/TR/css3-writing-modes/>, (accessed 2011-01-14).

## CA1742

## 動向レビュー

## ライブラリー・グッズの可能性—ミュージアム、米・英の国立図書館の事例を通して

## 1. はじめに：日本におけるライブラリー・グッズの歴史と現状

このところ、大学図書館を中心にライブラリー・グッズへの関心が高まっている<sup>(1)</sup>。図書館における広報活動へ関心が寄せられると同時に、広報の一手段としてのグッズも徐々に注目を集めるようになってきたということが一因として考えられる。しかしながら、米国及び英国を中心として、図書館内の店舗でのグッズ販売やオンライン・ショップへの展開など活発な動きが見られるのに対し、日本の図書館でのグッズ制作・販売は、比較にならないほど小規模にとどまっている。

日本でのライブラリー・グッズへの関心は、少なくとも1980年代に遡ることができる。個々の図書館で行われてきたグッズ制作の歴史を遡るのは実質的には不可能であるが、比較的早い例としては、1983年、図書館サービス研究分科会（私立大学図書館協会東海地区部会研究部企画広報研究分科会の前身）の広報グループで始められたグッズの共同制作が挙げられる<sup>(2)</sup>。この共同制作は、複数の図書館で同一デザインのグッズを制作し、コストを抑えることを目的としていた。1983年に呼びかけを始め、翌年15館でポスターを作成したのが第一回となる。以降も同グループでの共同制作は続けられ、年ごとに掲示用紙、本のしおりやブックカバーを作成した。1992年度には日本図書館協会（JLA）の協賛へ、そして1995年度にはJLA主催の事業へと移管され、同年度には私立大学図書館協会、全国公共図書館協議会、専門図書館協議会、全国学校図書館協議会などが協賛し、紀伊國屋書店が後援となるが、1999年度を最後に、JLAの主催による共同制作は終了した。2005年度、2007年度には前述の分科会による共同制作が行われた<sup>(3)</sup>が、継続的な事業とはなっていない。

個々の館の活動としては、大谷大学図書館<sup>(4)</sup>、早稲田大学図書館<sup>(5)</sup>で、無料でダウンロードできる建物の写真や蔵書の画像を利用したブックカバーやカレンダーなどを用意しているほか、横浜市立図書館が2009年に横浜開港150周年を記念し、期間限定で同じく無料でダウンロードできるオリジナルブックカバーを作成していた例が挙げられる<sup>(6)</sup>。

有料販売については事例が少なく、東北大学附属図書館が2007年に東北大学創立100周年を記念して、同館の漱石文庫にちなんだ羊羹「漱石の愉しみ」を販