

The Electronic Text Center
at the
University of Virginia Library
バージニア大学図書館における電子テキストセンター

David Seaman
Director, Electronic Text Center
Alderman Library
University of Virginia
Charlottesville, VA 22903, U.S.A.
Tel.: 804-924-3230 Fax: 804-924-1431
E-mail: etext@virginia.edu
URL: <http://etext.lib.virginia.edu>

Kendon Stubbs
Associate University Librarian
Alderman Library
University of Virginia
Charlottesville, VA 22903, U.S.A.
804-924-3026 Fax: 804-924-1431
E-mail: kstubbs@virginia.edu
(訳: 杉本重雄・図書館情報大学)

概要

電子テキストセンターは図書館における人文科学分野向けの計算機を利用したサービスであり、次の2点
がその中心的な役割である。

- 電子テキストとイメージのオンラインアーカイブを構築し維持すること。
- 人文科学分野の広い範囲の利用者コミュニティを築き、維持していくこと。

5年間の活動を通して、オンライン利用可能ないろいろな言語のSGMLテキストを多く集積してきた。また、数多くの講習会を通じてトレーニングを受けた多くの学生と教官が、我々のサービスを教育と研究の場で利用する高度な利用者となっている。

キーワード

電子テキストセンター、電子テキスト、SGML、HTML、デジタルイメージ、全文データベース、デジタル図書館、オンライン・ライブラリ、人文科学における計算機の利用、インターネットのトレーニング、教育と技術

はじめに

5年近く業務を行ってきたバージニア大学図書館・電子テキストセンターはオンライン・コレクションの蓄積を進めるとともに大学内外での電子テキストの利用の促進に努めている。本電子テキストセンターは、Oxford English Dictionary、English Poetry Database、Patrologia Latinaといった1990年代始めに出版された人文科学分野に関連する商用の電子テキストの最初の波に対応するため図書館によって設立されたものである。1992年にバージニア大学図書館では、利用者からの要望の高まりに先立ち、人文科学における計算機の利用の促進と発展を担うセンターのために場所と機器、ならびにスタッフを導入することを決めた。このセンターの仕事は、電子テキストの作成と分析のための機器を利用可能にすること、そうした機器や技術のトレーニングを行なうこと、バージニアの人文科学におけるSGMLの利用の発展を担う活動の中心として働くこと、インターネットを介しては提供されていない電子テキストの利用のための場所を提供することである。

所蔵電子テキスト

現在、インターネットを介して利用できる数千の電子テキストとイメージを所蔵している。以下に示すように、その内には商用のものが含まれ、それらはバージニア大学あるいはVIVA (Virtual Library of Virginia、訳者注)の39のサイトでのみ利用することができる。UNIXのサーバにロードし、インターネットを介して利用者に提供するため、我々はこうした大きな商用の電子テキストファイルをCD-ROMではなくSGMLのデータファイルとして蓄積してきた。

The Oxford English Dictionary (25 volumes)
English Poetry Database (4,500 works)
English Verse Drama Database (1,500 titles)
English Prose Drama Database (1,500 titles)
American Poetry Database (4,500 works)
African-American Poetry Database (2,500 poems)

以下のものの利用はバージニア大学内のみに限られている。

The Patrologia Latina (221 volumes)
The Old English Corpus (3,000 items)
British philosophy: 1600-1900
American Civil War newspapers

(訳者注：バージニア州の高等教育機関における情報資源の効率的な利用のために州内の学術図書館で構成するコンソーシアム)

このほか文学、歴史、哲学、および宗教関係の資料が数多くあり、またいろいろな言語で書かれている。その中には18・19世紀に出版された英語の文学書と歴史書(挿し絵のあるものも多く含まれる)、フランス語、ドイツ語、ラテン語、日本語および中世英語のものがある。後に示したものの大半はインターネット利用者に開放されている。現在、電子テキストセンターのWebサーバには世界中から一月当たり100万件のアクセスがある。これを1日当たりに換算すると、4000の異なるマシンから、15,000ページの資料が33,000回アクセスされたことになる。

我々の電子テキストは全てSGML(Standard Generalized Markup Language)を用いて記述されており、World Wide Webを介した利用のためにそれらは自動的にHTMLに変換される。我々が作成ないしマー

クアップした電子テキストは Text Encoding Initiative (TEI) のガイドラインに従い、センターのスタッフによってタグ付けされている。また、新しく電子テキストを作成するためにバージニア大学および他大学の学生や教員と協力する機会が増えてきている。Japanese Text Initiative は新たな電子テキストを作成するために協力した好例であろう。これはバージニア大学の Kendon Stubbs とピッツバーグ大学の野口幸生が共同で作成したもので、SGML で記述され、インターネット上で検索が可能な日本文学資料のコレクションであり、現在も増えつつある。

英語の電子テキストはバージニア大学図書館のスペシャルコレクションから選ばれたもので、多数の Jefferson の手紙、Mark Twain 関係の資料、19 世紀のアフリカ系アメリカ人に関する歴史資料が含まれている。教員も学生も電子テキストセンターを使うにつれて、目的に合わせて我々の資料を整形したり補足したりしたテキストを作るようになり、オンライン情報の消費者から生産者になってきている。19 世紀アメリカの小説家に関する教材である Stephen Railton 教授の Mark Twain in his Times は本学教員によって電子的な資料が利用された好例である。

オフライン利用のための電子テキスト

我々の所蔵資料の多くは WWW を介して利用可能であるが、法律のもしくは技術的な理由でオンライン利用可能にできないものもある。それらは、Global Jewish Database、Thesaurus Linguae Graecae (古代ギリシャの作品 8000 点)、Perseus (ギリシャ語テキストとイメージからできたハイパーテキスト)、CETEDOC (ラテン語の神学資料)、Admyte (中世スペイン語)、Robert Musil と Immanuel Kant および Thomas Aquinas の作品、19 世紀アメリカの詩集である。

アクセス

電子テキストの多くがオンライン利用可能であるので、我々は単一の検索ソフトウェア (OpenText 社製) でコレクションの検索を行い、Netscape のような WWW ブラウザを共通でかつ親しみやすいユーザインタフェースとして利用している。我々は、利用者による検索とブラウジングのために、SGML から HTML への変換等のための CGI スクリプトを開発してきた。これらによってもたらされる利用者のメリットは明らかである。すなわち、一つのデータベースの使い方さえ知れば、任意の所蔵資料の検索ができるようになり、CD-ROM の場合のようにディスク毎に使い方が異なるといったフラストレーションを感じるようになる。

利用者

本センターの主要な目的はバージニアにおける人文科学分野の電子的資料の幅広いユーザ・コミュニティを築くことである。この目的のため、我々は定期的な講習会を開いている。そこでは資料のスキャニング、HTML、電子テキストやイメージの利用方法や生成方法に関する様々な話題について講習している。この 5 年間、作文の講義を受講している学部 1 年生から、アングロサクソン文学、アメリカ学、ラビ回答書、中世フランス語、その他様々な教育研究プロジェクトに携わっている大学院生に至るまで個別の利用者に対する仕事を毎日行ってきた。また、バージニア大学の Institute of Advanced Technology in the Humanities (訳者注：人文科学において利用できる先端技術の研究開発を行っている) の人達を支援してこれたことを喜びとしている。

大学出版局、学術出版社、学術雑誌の制作者、ならびに拡大しつつあるデジタル図書館との連携を強めるよう努力している。我々は、The Visual Anthropology Review と Essays in History というバージニア大学

で編集・制作されている2つの学術雑誌を支援し、我々のサーバ上にオンライン出版のための場所を提供している。また、我々は、Univeristy Press of Virginia から出版された Timothy D. Pyatt による Guide to African-American Documentary Resources in North Carolina (Charlottesville, 1996) と Michael Plunkett による Afro-American Sources in Virginia の検索可能な電子テキストを作成した。現在、雑誌 Studies in Bibliography の50年分の SGML への変換を終了しつつある。これはインターネット上での提供を目指したもので、バージニア大学の書誌学会 (Bibliographical Society) が進めている大きな出版プログラムの一部として行われたものである。

さらに、The American Heritage Virtual Archive Project (EAD) と The Electronic Archive of Early American Fiction という二つの大きなプロジェクトを助成を受けて進めている。後者は本会議におけるもう一つの論文で述べている。前者は National Endowment for the Humanities (NEH) からバージニア、Duke、Stanford、カリフォルニア大学バークレー校 (UCB) に対して与えられたもので EAD(Encoded Archival Description) タグを用いて大量のアーカイブ検索支援用資料を SGML 化するものである。

本電子テキストセンターは同様な試みを始めようとする機関にとってのモデルとなることが多くなってきている。これまでに数多くの図書館員や研究者が本センターを訪れ、それらには Harvard、Duke、Indiana、Johns Hopkins、Iowa、Yale、Columbia、Chicago、Kentucky、UCB、Virginia Tech.、Richmond、UNC Chapel Hill (ノースカロライナ大学)、UT Austin (テキサス大学)、Emory の各大学と National Humanities Center、ヨーロッパから英国図書館、Oxford、Cambridge、Nottingham、Glasgow、Leiden、Bielfeld、Groningen、日本の国立国会図書館、大阪外国語大学、図書館情報大学、オーストラリアの Sydney、Macquarie、それに Curtin 大学が含まれている。こうした交流は電子テキストとイメージに関するサービスを各地に広げることを推し進めることになるので、我々にとって重要な活動である。

今後の5年間は、人文科学研究者にとって VRML(Virtual Reality Modeling Language) やインターネット全体を対象とする検索ツールがさらに利用しやすくなることで、これまでの5年間と同様に非常に大きな変化を目にすることになるであろうと考えている。我々のユーザ・コミュニティの要望もニーズも徐々に洗練されたものなりつつあり、出版社が作り出す電子テキストとイメージの量は飛躍的に大きくなりつつある。激しい技術革新の中でも長期に渡って生き残るようにデータは SGML のような標準化された形式で蓄積されねばならず、また長期に渡ってオンラインサービスを存続させるには図書館と図書館が持つ技術が中心的役割を果たさなければならない。我々のサービスにおけるこうした基本的認識は利用者や出版社が進歩しても変化することはない。

参考文献

The Modern English Collection <http://etext.lib.virginia.edu/modeng/modeng0.browse.html>

British Poetry 1780-1910 <http://etext.lib.virginia.edu/britpo.html>

Mark Twain in his Times <http://etext.virginia.edu/railton/>

Japanese Text Initiative <http://etext.lib.virginia.edu/japanese/>

Electronic texts from Special Collections <http://etext.lib.virginia.edu/speccol.html>

Timothy D. Pyatt. Guide to African-American Documentary Resources in North Carolina. Charlottesville, 1996. <http://www.upress.virginia.edu/epub/pyatt/>

Michael Plunkett. Afro-American Sources in Virginia. A Guide to Manuscripts. Charlottesville, 1995. <http://www.upress.virginia.edu/plunkett/>

Essays in History <http://etext.lib.virginia.edu/journals/eh/>

The Visual Anthropology Review <http://etext.lib.virginia.edu/VAR/>

SGML Resources <http://etext.lib.virginia.edu/sgml.html>

VIVA, the Virtual Library of Virginia <http://www.viva.lib.va.us/>

The Bibliographical Society of the University of Virginia <http://etext.lib.virginia.edu/bsuva/>

The Electronic Archive of Early American Fiction <http://etext.lib.virginia.edu/eaf/>

The American Heritage Virtual Archive Project (EAD) <http://etext.lib.virginia.edu/ead/>