

Journal of Japan Society of Information and Knowledge

# 情報知識学会誌

Vol.12 No.4 (Jan. 2003)

~~~~~ 目次 ~~~~~

## 特集「XMLによる情報知識システム」

|      |                                                               |       |                                           |    |
|------|---------------------------------------------------------------|-------|-------------------------------------------|----|
| 巻頭言  | .....                                                         | 石塚 英弘 | 1                                         |    |
| 論文   | 日本の記録史料記述 EAD/XML 化と記録史料管理—記録史料管理過程における EAD 利用の位置をめぐって—       | ..... | 五島 敏芳                                     | 3  |
| 論文   | バイオ研究者間の情報共有のためのプロテオーム XML 記述方式                               | ..... | 水口 弘紀、上條 憲一、次田 皓                          | 22 |
| 論文   | 健診情報のための電子的交換規約                                               | ..... | 原 正一郎、杉森 裕樹、古海 勝彦、東福寺 幾夫、窪寺 健、河合 正樹、吉田 勝美 | 32 |
| 論文   | 「XMLに基づくコンピュータ処理可能な構造化デジタルオブジェクト」の概念による Web リソースの検索システムの研究と開発 | ..... | 李 穎、石塚英弘                                  | 53 |
| 論談   | 「ドキュメント・マネジメント」による知識創発型の地域社会づくり                               | ..... | 西村 健                                      | 69 |
| 解説   | 伝統的知識と知的所有権 — 解説                                              | ..... | 名和 小太郎                                    | 77 |
| 追悼   | 長瀬真理先生の訃報に接して                                                 | ..... | 根岸 正光                                     | 84 |
| 再掲載  | 新しい編集規程が整備・制定されました                                            | ..... | 安永 尚志                                     | 85 |
| お知らせ | 論文賞創設                                                         | ..... | 安永 尚志                                     | 96 |
|      | 第 11 回研究報告会 発表論文募集                                            | ..... | 国沢 隆                                      | 97 |
|      | 特集「科学技術データの活用」の論文公募                                           | ..... | 国沢 隆、菅原 秀明、宇陀 則彦                          | 98 |
|      | 情報知識関連新刊図書一覧                                                  | ..... | 平田 周                                      | 99 |



情報知識学会

## 特集「XMLによる情報知識システム」に寄せて

担当編集委員 石塚 英弘

情報知識システムは情報知識学会(以下、本学会)にとって最も重要なテーマの一つである。なお、ここで情報知識システムとは、情報ないし知識を扱うシステムあるいは系のことを言う。また、現在、電子商取引への適用が進み、注目を集めているXML(Extensible Markup Language)は、その規格が策定された当初は、ドキュメントを含む様々な情報を構造化して表現しうる枠組みであり、情報の交換フォーマットにも適用できると考えられていた。その点ではXMLは情報知識を扱うシステムあるいは系に相応しい基盤技術と言える。

また、XMLは本学会とも関係が深い。たとえば、本学会は、XMLの前身の規格であり、HTML(HyperText Markup Language)の規格の元でもあるSGML(Standard Generalized Markup Language)に学会設立当初から注目し、SGMLを用いた学会誌の出版に日本で初めて成功した。また、SGMLないしXMLの技術と適用例を紹介しあう会:「SGML研修フォーラム」、最近の名称は「SGML/XML研修フォーラム」を毎年開催するなど、XMLの啓蒙普及にも努めてきた。

そのような経緯もあり、特集「XMLによる情報知識システム」を本号(12巻4号)に組むことになった。

標記の特集テーマへの応募原稿は通常の審査の手順、すなわち、編集委員会による査読者の選定、研究論文の場合は2名の査読者、論談の場合は1名の査読者による査読、査読報告書の編集委員会での審議と採否の決定、により処理された。その結果採択された研究論文4篇、論談1篇を本特集に掲載した。なお、著者の中に編集委員が含まれる場合は、採否の公正を保証するため、その編集委員は査読者の選定、査読、論文の採否などに関与しな

いとの方針を立て、それを実行した。

XMLは様々な分野に適用可能であると言われているが、今回の特集の掲載論文・論談の対象分野は、歴史学、生物学、医学、医薬品情報、電子政府であり、図らずもXMLの適用範囲の幅の広さを実証した結果となった。また、これら幅広い対象分野を情報と知識の観点から横断的に見ることが本学会の特徴であり、その点からも今回の特集テーマは本学会に相応しいと感じた次第である。

以下、査読を経て採録された研究論文と論談について、掲載順にその内容を簡単に紹介したい。なお、著者の敬称は省略した。

五島の論文「日本の記録史料記述 EAD/XML化と記録史料管理」は、アメリカ記録史料管理学会がアメリカ議会図書館の協力を得て策定した記録史料記述の国際規格:EAD/XML(Encoded Archival DescriptionのXML版)の日本語版移植に関する最初の本格的な実証論文である。なお、五島は日本近世を専門分野とする歴史研究者である。本論文はこれまであまり論じられなかった記録史料管理におけるEADの利用について、具体的かつ実践的に論じており、人文科学とりわけ歴史学分野においては有用な論文と言えよう。

水口らの論文「バイオ研究者間の情報共有のためのプロテオームXML記述方式」は、当該分野において最近注目されているプロテオーム研究について専門分野以外の研究者にも理解できるように解説した上で、プロテオーム解析実験を対象とするXML文書を定義し、データ共有のためのXML記述方式を提案している。この記述方式では実験条件も記述可能なことが特長である。なお、2DゲルデータのためのXML記述方式は世界最初である。また、そのエディタおよび文書を登録検索可能なサーバも開発した。実験に参加しなかつ

た研究者もデータを利用できるようになるため、本システムはプロテオーム研究の促進に寄与するであろう。

原らの論文「健診情報のための電子的交換規約－ Health Data Markup Language (HDML) －」は、健康診断データを異なるシステム間で交換、保存するためのXMLによる交換フォーマットを提案し、その規約概要を述べ、実証試験を行った結果を評価したものである。健診データの有効性を高めるためには、効率的な施設間の電子的データ交換を実現することが望ましい。そのため本法により実現される健診データ交換標準化の有用性は極めて高いと考えられる。なお、この規約は日本総合健診医学会情報委員会と日本保健福祉医療情報システム工業会との合同委員会により策定されたとのことである。なお、本規約は、W3CのXML関連規格の一つである携帯端末用のHDML (Handheld Device Markup Language) とは別の規格である。

李らの論文「『XMLに基づくコンピュータ処理可能な構造化デジタルオブジェクト』の概念によるWebリソースの検索システムの研究と開発」では、学術情報の検索機能と精度の向上のために、XMLに基づくコンピュータ処理可能な構造化デジタルオブジェクト、すなわち、CPDO (Computer-Processable Digital Object) と称する情報の最小単位を構成要素とする構造化情報表現を考案した。これにより、複雑な構造を持つ学術文献などが適切に構造化される。また、XML schemaを用いて表現されるため、汎用的である。次いで、その概念に基づいて、典型的なファクトデータベ

スであるMerck Indexのデータを例として、Webリソースの検索システムのプロトタイプを構築し、この方法の有効性を示している。

西村の論談「『ドキュメント・マネジメント』による知識創発型の地域社会づくり」は、自治体における電子政府では、従来行われている縦型、すなわち、政府指定の事務の請負執行から、横型、すなわち、地域を横にネットワーク化して地域において発生する複合課題に対処する方式に転換すべきであるとし、それにはXMLに基づく「ドキュメント・マネジメント」のプラットフォーム構築が重要であると主張している。長年に亙る自治体での実務経験およびその業務システムの開発経験を持つ著者ならではの情報が、本論談には含まれている。電子政府実現のためのe-Japan計画の中で「知識創発型社会づくり」が提唱されている現在、本論談は大いに有用と思われる。

本特集は、全体を見通して指導された安永編集委員長、興味深い研究を見出して投稿を勧めてくださった編集委員、投稿者の皆様、多忙の中で時間を割いて査読してくださった査読者の方々、年末年始の休みを削って作業してくださった事務局と編集印刷所の方々など関係各位のご尽力の賜物である。心より御礼申し上げます。

なお、本号には特集以外にも解説など貴重な原稿をいただいた。故長瀬理事への追悼文には心を打たれた。長瀬理事はテキストデータベースにSGML/XMLの技術を導入され、国際的に活躍されていた。お元気であれば、本特集に論文をいただけたに違いない。合掌。

日本の記録史料記述 EAD/XML 化と記録史料管理  
 —記録史料管理過程における EAD 利用の位置をめぐって—  
 An Application of EAD/XML to the Japanese Archival  
 Descriptions and the Management of Archives:  
 The Position of EAD, Encoded Archival Description  
 in the Management Process of Archives

五島 敏芳 \*

Haruyoshi GOTOH

記録史料記述の国際標準の一つ「国際標準：記録史料記述の一般原則；ISAD(G)」を、日本の記録史料記述、具体的には記録史料目録へ適用する試みは多く存在する。しかし、その記述構造を反映した電子化例は決して多くない。筆者は、ISAD(G) 準拠の記録史料記述 XML 化の経験と先行の EAD 化実践例をふまえ、記録史料記述電子化の事実上の国際規格 EAD, Encoded Archival Description (符号化記録史料記述；SGML サブセット) を採用し適用実験をおこなった。その結果、適用可能であることが確認でき、一定程度の困難が残るものの実用上の問題を克服しうることが判明した。この論文では、記録史料管理のいくつかの場面で発生する記録史料(記述)情報と、EAD 化に必要な情報とを、なるべく調和させる形で記録史料記述(記録史料目録)を EAD 化するための考え方や手順を紹介したい。これにより、記録史料管理過程における EAD の位置を明らかにする。

After the adoption of ISAD(G), *General International Standard Archival Description*, it was reviewed immediately in Japan. There are many experiments on the application of ISAD(G) to the Japanese archival description (or archival finding aids), whereas there is no application experiment of XML for the Japanese archival descriptions which applied ISAD(G) except for my experiment. However, in my experiment, I did not use EAD, *Encoded Archival Description* (SGML subset) which was a de-facto international rule for the electronic archival description, because I thought that EAD was complex and difficult, and it had the problem on practical use in Japanese archives.

This time, I examined some examples of practicing encoding archival finding aids. As the result of examination, it clarified the possibility of solution of the problem on practical use, though some unsettled problems remain. This article shows an idea and a procedure of the creation of EAD-encoded archival descriptions (or finding aids). I would like to point out that EAD can be used as a means to integrate various archival information in the management process of archives. In conclusion, this article clarifies the position of the use of EAD in the management process of archives. I hope that this article contributes to the spread of encoding archival descriptions (finding aids) in EAD on Japanese archives.

キーワード：EAD/XML, 記録史料記述 archival description,  
 適用実験 application experiment, 記録史料管理 management of archives,  
 記録史料目録 archival finding aids

\* 国文学研究資料館史料館  
 E-mail:gotoh@nijl.ac.jp

## 1. はじめに

最近 10 年の間、記録史料記述の国際標準が定められてきた。その一つに「国際標準：記録史料記述の一般原則；ISAD(G)」<sup>[1]</sup>があり、はやくから日本でも紹介されている。この ISAD(G) については、日本の記録史料記述への適用実験が既にいくつも存在する<sup>[2]</sup>。ただ、その国際標準の目的にうたわれる“記録史料に関する情報の検索・交換の便宜を図ること”の具体的な実現、とくに記録史料記述の電子化を考慮した ISAD(G) の適用実験または実践の事例は、けっして多くはない<sup>[3]</sup>。

筆者は、ISAD(G) 準拠の記録史料記述を単に電子化するのではなく、マークアップ言語 XML を利用し記録史料記述の構造を電子データとしての構造にも反映させる適用実験をおこなったことがある<sup>[4]</sup>。しかし既に、記録史料記述の構造を電子データの構造へ反映できる事実上の国際規格が存在していた。XML の母語 SGML のサブセット、EAD, *Encoded Archival Description* (符号化記録史料記述) である<sup>[5]</sup>。EAD は、近年日本でもよく紹介され<sup>[6]</sup>、実は筆者も先の XML 利用の実験の際にその存在を知っていた。構造の反映と多様な利用可能性を確保するため、緻密に設計された EAD は、XML の浅薄な理解の水準にとどまっていた筆者にとって、当時、難解で複雑なものに思えた。その実感は、おそらく日本の記録史料管理に関わる多くの人びとにとっても、共感できるものと判断し、また EAD の日本語 (あるいは 2bytes 文字といってもよい) 対応の不十分さ等などをあげつらって利用を避けていた。

一方、記録史料記述への国際標準の適用が記録史料情報の検索・交換の便宜を目的とするならば、既存かつ事実上の国際規格の利用は、情報共有の可能性や効果の点で不可避であるだけでなく魅力的であった。そこで、EAD の内容と海外の EAD 適用事例の検討を進め、実際に日本の記録史料記述へ EAD を適用してみた。結果、記録史料記述の EAD 化は、その情報要素の基本的な部分のみの利用

からはじめれば、十分実現可能であることが明らかになった。記録史料記述の構造を反映した、テキストファイル形式の電子データの作成のみならず、EAD を XML のなかで扱い、海外の既存の EAD 化事例における表現を模倣することで、ウェブ・ブラウザ上での表現もひとつおき実現できた。

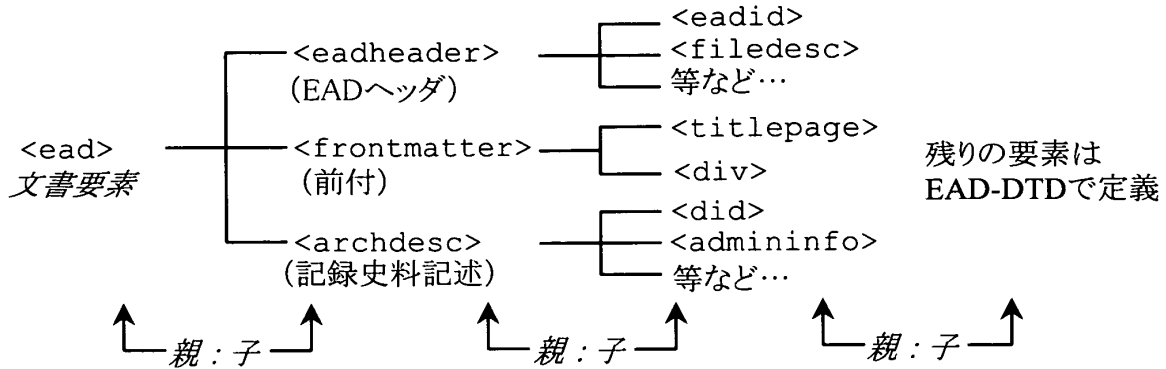
以下その内容を紹介していく。なお、この報告の中で紹介される日本の記録史料記述に対する EAD 適用実験は、おそらく日本で最初の事例であろう (2002 年 5 月当時) ことも付言しておく<sup>[7]</sup>。

## 2. EAD の基本的な構成

EAD が SGML のサブセットでありマークアップ言語であることは前述したが、その性質により当然、情報要素をくくる「タグ」(および閉じタグ)が存在する。それら EAD のタグは、記録史料記述の記述要素に対応し、ISAD(G) だけでなく、書誌情報の標準枠として定着している MARC, *Machine-Readable Cataloging* や、メタデータ規則の一つ Dublin Core の、各要素とも対応している<sup>[8]</sup>。

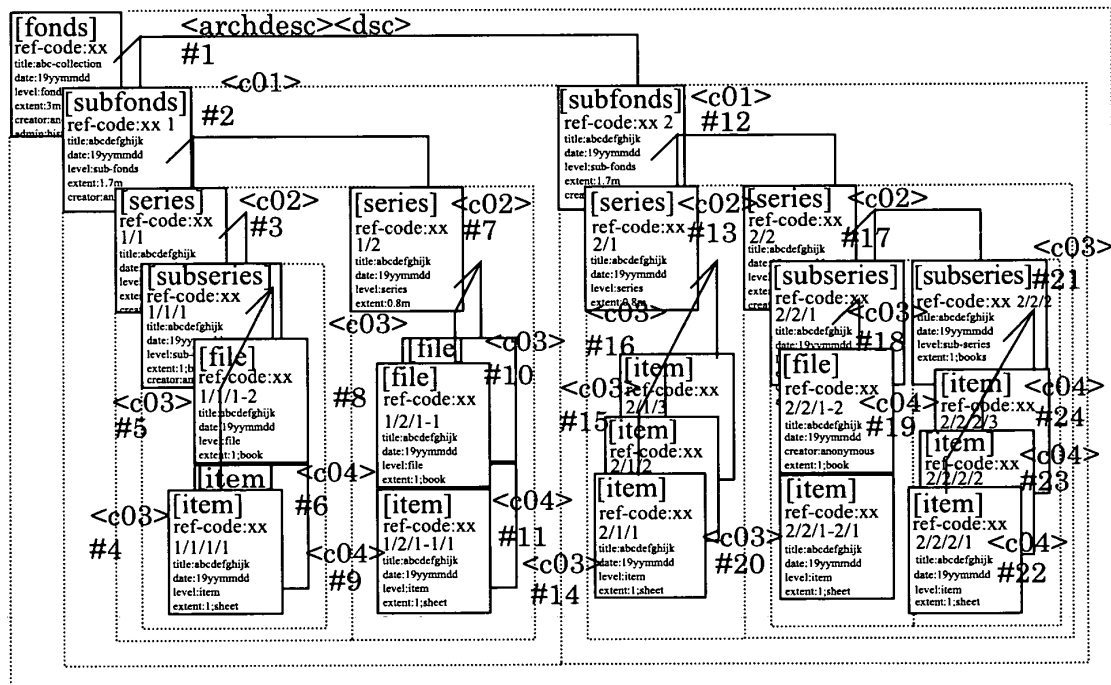
EAD のタグは、およそ図 1 のように構成され、それぞれ階層的な構造を持つ。大きく EAD ヘッダ・前付け・記録史料記述の三つの部分に分かれ、とくに記録史料記述の部分に中心となるデータを収める。その主要なタグについては、EAD アプリケーション・ガイドライン付録 A「最小の推奨検索手段要素」に示されている。

EAD の記録史料記述部分は、記録史料の群や小群からなる階層的な秩序を反映できるようになっている<sup>[9]</sup>。そこで利用可能なタグの設定からは ISAD(G) のマルチレベル記述をおよそ実現しているものの、フォンド・レベル記述とサブ・フォンド・レベル以下の記述が等質な関係にないことを指摘できる。フォンド・レベル記述に対応するであろう <archdesc>タグ中に、サブ・フォンド・レベル以下の各記述を <dsc>タグ下に収め、



※アプリケーション・ガイドライン Figure 6.2.1b. 階層的な樹系構造としてのEADより

図1: EAD ファイル内の構成 (概要)



※「#」を付した数字は、EADファイル内での当該データ配置の絶対順を示す。

図2: EAD ファイル内の構成 (概要)

さらにサブ・フォンド・レベル以下の各レベルごとに<c01>~<c12>まで入れ子状に記述データを格納する。つまり、フォンド・レベル記述以下のレベルの記述を、フォンド・レベル記述に内包させている(図2参照)。記録史料が、基本的にフォンドを単位として把握されることを考えれば、合理的な発想である。また、入れ子状の記述データ格納は、各記述レベルのデータの配置を決定し

ている。整列秩序を与えなければ無秩序に存在するデータベースのレコードのように、異なった記述レベルのデータが混在する、ということではなく、ファイル内に記録史料の内的な階層構造にそくしたシーケンシャルな配置を可能にしている。

なお、EADの記録史料記述部分のタグに設定されている階層レベルと、記述レベルとは、一致していても一致しなくてもよい

ようである。たとえば、<c02>のレベルはつねにサブ-フォント記述でなければならない、ということはない。研究の進展により今後、記録史料記述が改変される可能性がある場合、記述レベルとタグの階層レベルとの対応を固定しないほうがよいことであろう。

記録史料記述の EAD 化に際して、とくにウェブ・ブラウザ上での表現を想定し XML を利用する場合、一定のファイル構成が求められる。まず、EAD の設計図ともいえる次のようなファイル群：ead.dtd, eadbase.ent, eadchars.ent, eadnotat.ent, eadgrp.dtd, eadsgml.dcl, xmlchars.ent は必須である。もちろん、これら DTD 類はなくとも、EAD のタグを熟知していれば、そのタグをもとに XML を書くことはできようが、EAD の文法に正確かどうかは、これらの DTD 類なくして確認できないだろう。そして当然、EAD のタグを付された記録史料記述の XML ファイルは必要である。EAD 化した記録史料記述のテキストファイルを XML ファイルとするには、ファイル冒頭に XML 宣言を加えるだけでよい。たとえば「<?xml version="1.0" encoding="shift-jis"?>」とすれば、日本語・Shift-JIS の環境で利用できる。続いて文書型宣言に、EAD の DTD 類の引用とローカルでの所在を記す。これらに加えて、XML ファイルを具体的にウェブ・ブラウザ等へ出力するため、スタイルシートが必要になる。単に内容を表示するだけであれば CSS で十分だが、表示オブジェクトへの変換だけでなくデータの構造の改変をとまなうならば XSL を利用しなければならない。なお、XSL に対応していないウェブ・ブラウザもある<sup>[10]</sup>。今回、筆者は既存の EAD 用の XSL を模倣したに過ぎないが、本来 EAD 表現用 CSS 例をも提示する必要がある<sup>[11]</sup>。そもそもウェブ・ブラウザでの表示を EAD 表現の確認手段とする場合、EAD 内の情報要素の表示のされ方自体が検討されなければならない。

EAD のタグは、ある動作や効果が期待さ

れる機能の側面から見ると、記述構造用のタグ、表現用（書式）タグ、ファイル内外の参照指示用のタグ等に、大きく分けることができそうである。こういった期待される機能は、XML の場合、実は前述の XSL なくして実現できない。記録史料記述の構造的範囲を明示するようなレイアウトや、表組み、強調体や斜字体といった文字の効果、関係資料へのハイパーリンク、等などのためのタグや属性が存在し、タグの組み合わせ方やデータの入力ルールといった仕様も設定されている。各機能を実現するスタイルシート類は、それらの仕様にもとづいて定められている。

すべてのタグを使用するわけではなく、すべての動作・効果の機能を期待するわけではない、とすれば、限定的な機能のスタイルシート類で十分である。そもそも XML を利用するという所与の条件からすれば、恣意的なタグの利用や解釈は発生しうる（後述）。

EAD の基本的な構成の確認からは、EAD を XML を通して利用する場合、最低限必要なファイルのセット、EAD ファイル内の最低限必要なタグ構成、が明らかになった。次に、既存・先行の EAD を適用した記録史料記述の例を確認する。

### 3. EAD を適用した記録史料記述例

ここでは、インターネットを通して入手した海外（アメリカ）の EAD 記録史料記述の XML ファイルを検討する（図 3 参照）。

一つは、コーネル大学図書館の EAD/XML 検索手段プロジェクト<sup>[12]</sup>による、いくつかの EAD 記録史料記述の XML ファイルである。

(1) の「Howard S. Liddell Papers」はコーネル大学の心理学教授の書類で、ここにあげた表示例はブラウザからの印刷出力にもたえる体裁を持っている。このファイルに使用されている XSL のスタイルシートでは、記録史料記述の主要な記述要素ごとにリンクのアンカーが埋め込まれていて、フレームで目次の画面をともに表示させ、目的の記録史料記述

(1). Howard S. Liddell Papers, 1920-1967 (部分)

Guide to the Howard S. Liddell Papers, 1920-1967

Collection Number: 14-23-1466

Division of Rare and Manuscript Collections  
Cornell University Library

Contact Information: Division of Rare and Manuscript Collections  
28 East 4th Street Library  
Cornell University  
Ithaca, NY 14853  
607/255-3030  
Fax: (607) 255-9624  
rare@cornell.edu  
185.255.3030.1466

© 1994 Division of Rare and Manuscript Collections, Cornell University Library

DESCRIPTIVE SUMMARY

Title: Howard S. Liddell papers, 1920-1967  
Collection Number: 14-23-1466  
Creator: Howard S. Liddell, 1895-1962  
Quantity: 15 cubic ft., 1 tape recording  
Forms of Material: Correspondence, lecture and research notes, course materials, financial and subject files, manuscripts, printed materials, photographs, diagrams, glass slides and negatives, movie film, tape recording.  
Repository: Division of Rare and Manuscript Collections, Cornell University Library  
Abstract: Personal and professional papers and materials of Cornell University professor of psychology, founder of the Behavior Farm Laboratory.

BIOGRAPHICAL NOTE

Howard Scott Liddell, 1895-1962, was professor of psychology and psychobiology and founded and directed the Behavior Farm Laboratory at Cornell University. Following Liddell's death, the lab was renamed the Liddell Laboratory of Comparative and Physiological Psychology.

CHRONOLOGY

INFORMATION FOR USERS

Cite As: Howard S. Liddell Papers, #14-23-1466. Division of Rare and Manuscript Collections, Cornell University Library.

SERIES LIST

|                                 |             |
|---------------------------------|-------------|
| Series I. Personal              | Box 1       |
| Series II. Correspondence       | Boxes 1-2   |
| Series III. Professional Papers | Box 3       |
| Conferences                     |             |
| Lectures                        |             |
| Students' Notes                 |             |
| Series IV. Research Papers      | Boxes 3-5   |
| Notebooks                       |             |
| Notes and Files                 |             |
| Research Files                  |             |
| Series V. Cornell Dept. Files   | Boxes 5-7   |
| Manuscripts                     |             |
| Financial                       |             |
| Subject Files                   |             |
| Series VI. Printed Material     | Boxes 7-9   |
| Series VII. Visual Material     | Boxes 10-20 |
| Photographs                     |             |
| Diagrams                        |             |
| Glass Slides and Negatives      |             |
| Movie Films                     |             |
| Series IX. Miscellaneous        |             |

CONTAINER LIST

| Description                                                            | Box or Drawer | Folder or Item No. |
|------------------------------------------------------------------------|---------------|--------------------|
| Series I. Personal                                                     |               |                    |
| Notebook: Prof. Simpson's lectures on the Endocrine Organs, March 1920 | 1             | 1                  |
| Exam Book: Notes on Endocrine Organs                                   | 1             | 2                  |
| Notebook: Prof. Simpson's lectures, March 1920                         | 1             | 3                  |
| Notebook: Physiology of Nervous System, Vol. 2                         | 1             | 4                  |
| Liddell Notebook: Series 1922, Learning Tests                          | 1             | 5                  |
| Notebook: Organic Chemistry                                            | 1             | 6                  |
| Notebooks: Physiology                                                  | 1             | 7-9                |
| Notebook: Physiology notes, nervous system                             | 1             | 10                 |
| Notebook                                                               | 1             | 11                 |
| Liddell accomplishments (list format) and biography                    | 1             | 12                 |

(2). Ezra Cornell papers, 1746-1888 (それぞれ部分)

Series VII: Estate Records

Series Description

Papers pertain to Ezra Cornell's will and estate, and include mortgages, warranty deeds, tax papers and documents, title abstracts, land contracts, bonds, receipts, and other papers.

Series Content

| Description                                                                                 | Box / Folder |
|---------------------------------------------------------------------------------------------|--------------|
| Mortgages, warranty deeds, tax documents, title abstracts, land contracts, bonds, receipts. | 78-80        |

Series VIII: Scrapbooks, Broadside, Maps, Photographs, Clippings, Ephemera, and Genealogical Information

Series Description

Several forms of material document the chief vocations and enterprises of the life of Ezra Cornell. Newspaper clippings kept by Cornell and later by members of his family chiefly concern the telegraph industry and the founding of Cornell University. Other topics and issues featured in the clippings include the growth of American railroads, the Republican Party, local and national politics and government including Ezra Cornell's candidacy and service in the New York State Legislature, and real estate especially pertaining to the Wisconsin lands; also mining, scientific advancement, medicine, and other issues. Other material includes agricultural catalogues, maps, particularly of lines of the developing telegraph industry and Wisconsin lands; photographs of Cornell and his family, broadsides; Ezra Cornell's "epitaphing" book" (1823-1860) which include financial and arithmetic lessons and calculations; the Manual of the Common Council of New York; material from the New York State Constitutional Convention of 1867, calling cards, railroad passes, the New York State Agricultural Society's medal presented to Cornell for his cattle, Cornell's ceremonial wedding socks, and a volume celebrating the Cornell Public Library. Also, biographical and genealogical materials, and printed materials concerning Cornell University.

For genealogical information, see also Rev. John Cornell, *Genealogy of the Cornell Family: Being an Account of the Descendants of Thomas Cornell of Portsmouth, N.H.* (New York: T. A. Wright, 1902) - digital copy available.

Series Content

| Description                                | Box / Folder |
|--------------------------------------------|--------------|
| Clippings, 1847-1873.                      | 81           |
| Clippings, 1852-1874.                      | 82           |
| Ezra Cornell's Epitaphing Book, 1823-1860. | 83           |
| Manual of the Common Council of New York.  | 85           |

[Ezra Cornell Guide] | [Ezra Cornell Collection] | [CDL Page]

図 3: 既存の EAD/XML 記録史料記述の表示例



(3). American Revolution Collection (部分)

Overview amrev1776.xml

Related Material **American Revolution Collection**

Scope and Content  
American Revolution Collection of the Connecticut Historical Society

Organization

CHS  
The Connecticut Historical Society

Compiled by NHPRC Project Staff

History  
EAD conversion sponsored by grant funding from the National Historical Publications and Records Commission, Grant # 98-101

Administrative Information

Back  
CHS Home

Connecticut Historical Society, March, 1999  
1 Elizabeth Street Hartford, CT 06105

---

**Collection Overview**

Creator: Connecticut Historical Society  
Title: American Revolution Collection  
Date: 1776-1786  
Abstract: Collection consists of records of the Council of Safety, whose function it was to handle the day to day affairs of Connecticut's wartime government; records of the Commissary department which was responsible for providing food and equipment for the soldiers; records of the Committee of Paytable which paid for supplies and services; payroll records of Col. Sheldon's Light Dragoons, the commissary and paymaster records of the Third Regiment, Connecticut Line, orderly books, journals and correspondence describing activities of Connecticut persons during the Revolution.  
Extent: 11 boxes; 204 folders; 2191 items  
Location: Manuscript Stacks

---

**Related Material**

An index of catalog cards is available to access related material in other collections at the item level. Access is through writer, recipient and date. The card catalog is located in the library reading room.

---

**Scope and Content**

Collection consists of records of the Council of Safety, the Commissary, the Committee of Paytable, Sheldon's Light Dragoons, the Third Regiment, Connecticut Line and journals, orderly books and correspondence describing activities of Connecticut persons during the Revolution.

The records have been organized into nine series:

1. Records of the Council of Safety: 1774-1783
2. Commissary Records: 1774-1783
3. Pay Table Records: 1774-1783
4. Colonel Sheldon's Light Dragoons: 1770-1783
5. Third Regiment, Connecticut Line: 1770-1780
6. Account Books: 1775-1783
7. Naval Affairs: 1773-1789
8. Correspondence: 1774-1833

---

**History**

American colonists in the mid eighteenth century felt that the economic policies of Great Britain were unfair and that they did not have a sufficient voice in making decisions regarding these policies. Colonial boycotts of British goods did sufficient damage to force the repeal of the Stamp Act in 1766 and the Townshend Acts in 1770. When Parliament punished Massachusetts for destroying thousands of pounds of tea in Boston Harbor, the organized groups that had introduced these boycotts rebelled. The colonists had not intended to wage a war and were totally unprepared to provide soldiers with the weapons, training, clothing and food needs.

---

[Return to top of page.](#)

**Series 1: Records of the Council of Safety**

Title: Series 1: Records of the Council of Safety  
Date: [1774 - 1783]  
Extent: 1 box; 135 items  
Location: USCTH/AMREV/1776 - 1

Abstract: Series consists of muster rolls, returns, exemptions, substitutes, deserters, prisoners and soldiers missing in action.

Folder 0 Archives control file  
Folder 1A Muster Roll of Colonel Return Jonathan Meigs. 1779 (1 item)  
Folder 1B Muster Rolls of Captain Abel Pettibone. 1774 (3 items)  
Folder 1C Muster Rolls 1775 - 1781 (18 items)

(図 3 の続き)

(1). EAD 記述入力用 XML ファイル

2/6 ページ

```

http://www.nyu.edu/library/bobst/collections/findingaids/ </addressline>
</address>
<p>
<data>0 [date here] </date>
</p>
<address>
<addressline>Tarniment Library/Robert F. Wagner
Archives </addressline>
</address>
All rights reserved.
</p>
<publisher>New York University Libraries, Publisher</publisher>
</publicationstmt>
</filesdesc>
</profiledesc>
<creation>
Machine-readable finding aid derived from [original file format
and conversion before encoding].
<data>[date here] </date>
Machine-readable finding aid created by [Archivist Four].
</creation>
<language>
Description is in
<language>English </language>
</language>
</profiledesc>
<revisiondesc>
<change>
<data>[date here] </date>
<item>Electronic finding aid revised according to local
applications, by [Student Five]. </item>
</change>
</revisiondesc>
</archdesc>
<archdesc level="collection" lang="eng" type="inventory">
<did id="a1">
<head>Descriptive Summary</head>
<unittitle label="Title" encodinganalog="245bb00$a">
[Records of [title here]]
<unittitle>[data here, as per guidelines] </unittitle>
<unittitle>
<unittitle label="Accession number" encodinganalog="852bb\\$1">
repositorycode="NNU" >[[Tarniment or Wagner] [collection number]]
</unittitle>
<physdesc label="Quantity" encodinganalog="300bb\\$a">[extent
here] </physdesc>
<repository label="Repository">
<address>
<addressline>Tarniment Library/Robert F. Wagner
Archives </addressline>
</address>
</repository>
<origin label="Creator" >[corpname, persname, or famname
element here, with name of creator, as per guidelines]
</origin>
<physloc audience="internal" >[location here] </physloc>

```

(2). ISAD(G) 記述の EAD 変換フォーム

Archives Hub ISAD(G) / EAD Template

1/5 ページ

ISAD(G) / EAD Template

3.1: Identity Statement Area

Subfonds level record

3.1.1: Reference Code (Including NCA Repository Code)

3.1.2: Title

3.1.3: Dates of Creation Normalized Date

3.1.5: Extent of Unit of Description

3.2: Context Area

3.2.1: Name of Creator (Also add as Access Point manually)

3.2.2: Administrative/Biographical History

Repository Name (e.g. University of Nottingham)

Include Subfonds level records

Reset Sheet    Render to SGM    Render to ISAD(G)

Last modified Mon May 21 14:57:35 BST 2001

図 4: 記録史料記述の EAD 化のためのツール例

へジャンプさせることもでき、実際そういう提供もおこなっている。また、HTML形式に変換して提供している。

(2)は同大創立者の書類の場合で、豊富な情報を盛り込んであり、EADのさまざまなタグの利用を試験しているようである。記録史料記述のテキストとEADタグをあわせて、約768 Kbytesというファイル・サイズである。こちらはHTML版は用意されていない。かわりに、デジタル・オブジェクト——具体的には記録史料画像へのリンクが設定されており、リンクのポイントからは、下段に示したような当該記録史料の画像一覧が表示される。

いま一つは、コネティカット歴史協会のNHPRC EADプロジェクト<sup>[13]</sup>によるEAD/XML記録史料記述の成果である。

(3)の「American Revolution Collection」が一例で、こちらはもっぱらブラウザ画面上での利用のみを前提としているようである（もちろん印刷出力も可能である）。フレームではない一つの画面で、左側に目次を用意し、いくつかの重要な記録史料記述へのリンクのポイントを設けてあり、右側に実際の記録史料記述を配する。いくつか記述要素の区切りに、たとえば、あるシリーズ以下の記述が開始される直前に、きちんとトップ（表示内容冒頭）へ戻ることができるようにリンクのポイントがあり、閱讀上の配慮がなされている。

(1)～(3)のいずれもが、ISAD(G)の記述要素にあるような記録史料記述の主要部分について、どの記述レベルに対しても表示することを保証しているようだが、ファイルのなかに、もともとデータが入力されていない記述要素については仮にタグだけ存在していても表示されない。そのデザインには、およそ、フォンド・レベル記述からファイル（ないしアイテム）・レベル記述にいたる直前までの中間的な記述について文章で表示し、ファイル・レベル記述以下は請求記号・タイトル・年代などを1行に表示するリストの形式をとる、という大別がある。少なくとも(1)～(3)の機関で、これまでタイプライタ等での

打ち出しや刊行によって提供してきたような、記録史料目録——記録史料記述による検索手段の一つ——の紙面デザインと類似していることを推測できる。

これらは、ひとつの記録史料記述（記録史料目録）1ファイルとし、データ全てを示す形式で、基本的に表示内容の冒頭から画面をスクロールしながらの閲読を利用者に求めている。その利点は、記録史料ならではの特質——脈絡の把握なく1点の記録史料を正当に利用できない<sup>[14]</sup>という特質の理解に貢献することにある。

一方で、たとえば利用者が何らかの特定の事象に関わる記録史料すべてへのアクセスの情報を知りたいと思い、いくつかのキーワードを含むという条件に合致する記録史料記述すべての表示を要求したとき、その一覧と各記録史料記述の脈絡理解の助けとなる情報へ接近できる方法や道筋を表示するような仕組み＝検索システム、も求められる。こういったウェブ・データベースも実在しよう。

ただ、あえて(1)～(3)のような比較的基本的な実践例を紹介した理由は、多くの日本の記録史料保存利用機関等において高度な検索システムを実現できるほど電子化された記録史料情報の蓄積があるわけではなく、そもそも記録史料情報の電子化自体が記録史料管理過程のなかで位置づけられていない現状にあるためだ。まず基本的な実践に倣い、電子化作業を進め、電子化された記録史料情報を蓄積する必要がある。

付け加えておかななくてはならない重要なことは、(1)・(2)と、(3)の間の相違にみる、EAD適用の多様性である。これは、その表示上の個性の問題ではなく、その表示のあり方(XSLスタイルシートの内容)から判明する、EADタグの用法の相違である。たとえば、記録史料の内部構造を記述する<dsc>タグ内において、<cXX>タグを使ってデータ構造とともに階層のレベルを明示するか((1)・(2)の場合)、<c>タグを使って階層のレベルをデータ構造として明示せず属性値や実際の入れ子

状況から判断する（または表示等の際に XSL に判断させる）か（(3)の場合）、といった違いにあらわれている。

つまり、EADを導入しEAD/XML化を進めている記録史料管理組織（主体）によって、タグ用法や、記述要素および記述内容が変化してしまっている。この点は、当該の記録史料管理主体にとっての各記述内容の理解と関わっている。いっそうの記述内容自体の検討を進める必要がある。

さしあたり、一般的な記録史料目録ないし検索手段の表現において各記述レベルごとの記述要素が経験的に確定しており、各記録史料管理主体が継続的に提供してきた記録史料目録等の表現の枠組みが影響しているだろうこと、は指摘できよう。

記述内容認識の問題とは別に、EADタグの用法の多様性や自由度の高さには、いくつかの理由を推測できる。まず、EADファイル作成の技術的状況（環境）の影響である。欧米ではEADに限らずSGML系の文書の生成に特定のアプリケーションが普及しているように見受けられるが、EADファイル作成の実践はさまざまに試みられている<sup>[15]</sup>。図4には、(1)としてニューヨーク大学図書館で使用していると思われる入力用テンプレートを、(2)としてアーカイブズ・ハブというイギリスのプロジェクト<sup>[16]</sup>が提供していたISAD(G)記述からEADファイルを生成する（ただしフォンド～サブ・フォンド・レベルのみ）フォームを、それぞれ示した。(1)と(2)それぞれから得られるEADファイルの内容は、基本的な構造を同じくしながら詳細を異にしている。次に、既に電子化済みのデータ資産の活用を要求する圧力の存在である。それは、USMARCの情報要素とEADのタグが、緊密に対応していることからうかがえる。アメリカでは、MARCによって記録史料記述を電子化する動きが既に進展していたと思われ、これを無視することはできなかったのだろう。

既存のEAD適用例からは、具体的な表現と、EADタグの利用の自由度の高さを学ぶこ

とができ、既存の電子データ資産の活用に配慮されている可能性を見出すことができた。

#### 4. 日本の記録史料記述のEAD化

筆者が日本の記録史料記述に対して試みたEAD/XMLの適用実験は、筆者の勤務先での課題の影響を受けて、記録史料記述、なかでも記録史料の利用者へ提供する検索手段の一つとして（少なくとも日本では）重要な位置にある記録史料目録を、どのように利用者へ提示するか、EAD化した記録史料目録が有効だとすれば、それをいかに効率的に生成するか、という二つの問題関心から進められた。いきおい前者の問題にひきずられ、ウェブ・ブラウザ上での表示そして記録史料記述の効果的な表現について強く意識し、ISAD(G)の記述要素に対応するタグで記述のデータをくくり、ISAD(G)の持つ記録史料記述の表現可能性に対応するXSLのスタイルシートを作成すべく、XMLとXSLの両者のファイルを行き来する作業に大きな時間を割かれたが、その過程でEADのDTDの分析をおこなうことができた。EADへの理解をわずかながらも進めた結果、後者の問題については、実際の記録史料編成過程から生みだしたISAD(G)準拠の記録史料記述を単にEAD/XML化する実験の作業以外に、既に電子化されている記録史料目録のデータをEAD/XML化する作業を経ることで、一定の見通しを得ることができた。

この作業過程は、筆者の着想に大きな影響を与えたように思う。そこで作業の道程に沿って、以前作成したISAD(G)準拠のデータを使用したEAD/XML化の成果から、順に紹介していく。

図5には、(1)と(2)にEAD化した記録史料記述のXMLファイルをウェブ・ブラウザで表示（印刷出力）した際のスナップを載せた。このように表示するための、EAD記録史料記述のXMLソースが、(3)である。

一見してわかるように、表示に使用したXSL

のスタイルシートは、コーネル大学図書館からの借り物である。とはいえ、日本の記録史料目録にある情報要素に合わせた、なるべく違和感のない表示を実現するため、多くの改変を加えてある。また、記録史料画像へのリンクと表示のためには、ごく初歩的な CGI を作成して、それへのリンクを埋め込んでいる(図 5(1)の右側を参照)。スタイルシートに影響され、タグの用法は、コーネル大学図書館での実践に倣うことになった。

画面表示にはあらわれていないが、適用の試行錯誤の過程で、EAD のタグ・属性の用法にも、EAD のアプリケーション・ガイドラインから外れる改変を施した。ISAD(G) 準拠史料記述を XML ファイルへ具体化する際にも指摘した、日本の記録史料記述特有の問題が、その理由である。最たるものとして、年代(とくに和暦)記載と漢数字の処理をあげておく。そもそも EAD では、言語的な多様性は許容しているものの、地域的な表記慣習までは考慮されていない。そのため、たとえば<unitdate>タグに記録史料の年代表記を収める際、はじめ、原表記として登場し閏月を含む和暦を certainty というおそらく当該年代記載の確実性への評価を記す属性の値として格納した。のち、数値コードを収める normal 属性の存在に気づき、和暦を閏月にも考慮して西暦的に数値コード化した年代コードを同属性へ収め、和暦そのものは<unitdate>タグ内に記すことに改めた。これにより、たとえば<unitdate>の内容をキーにして、上位の<did>ノードの記述単位を年代順に並び替える、といった利用も可能になる。

問題は、和暦の特殊性・地域性を考慮した年代コードが、たとえば各地域の記録史料群の年代的分布の国際比較にたえる、標準的な内容と形式を持っているかどうか、である。さらには、そういった年代コードの生成や、多様な利用に際し和暦(および和暦の年代コード)と西暦の変換を、自動でおこなうサブ・システムの必要性、である。いずれも規格化と技術的な展開が解決しうるが、国際交換に

たえる最低限の共有可能な情報を用意する配慮は必要であろう。

一方で、こういった地域性を考慮したデータ生成の指針の提示や規格内容の改善を、EAD の策定主体へ要求することも必要である。今回の実験で EAD の DTD 改変(タグの新規増設)は可能だったが、あえて実行しなかった。上掲の例は他の地域でも起こりうる一般性を持つため、むしろ EAD 策定主体への何らかの要求のほうが妥当だと考えたためである。各地での EAD 適用時の不具合内容を蓄積し検討を進めなければならない。

今回判明した不具合は、ほかに次のようなこともあげられる。日本の一般的な記録史料目録では、文書の作成者だけでなく受取人も記される。この受取人という情報を、EAD のタグには反映できない。EAD に限らず ISAD(G)でも同様で、どうやら作成者と受取人をく差出・宛先〉のセットの概念としてとらえる思考がないようである。授受される文書は受取人の手許に残ることが当然で、受取人は記録史料群の集積主体＝出所と一致していることも当然である、という考え方が欧米では一般的なのかもしれない。文書が原本であるか写しであるか控えであるか、という区別を厳密に考えていないようである(この区別に関する EAD タグはとくに設けられていない)。日本の記録史料記述がアイテム・レベルを伝統的に重視してきたことは、記録史料を群としてのまとまりで認識する妨げとなっている一方で、古文書学との連携による緻密な研究の成果として世界に積極的に発信してよい内容のように思える。

受取人の記述要素は、実際には次のように処理した。利用者が目的の記録史料(記述)へ接近するためのキーワードを示すタグとして<controlaccess>があり、そのなかにも含む、というものである。同タグ下の、組織名を示すタグ<corpname>か、人名を示すタグ<persname>で、記述内容にそくして適宜タグを使い分け、記述内容の冒頭に「→」を加えることで区別できるようにした。あとは、デー

(1). 遠江国敷知郡舞坂宿文書 (部分)

(リンクされた画像)

遠江国敷知郡舞坂宿文書目録, 1604-1875

1/9 ページ

遠江国敷知郡舞坂宿文書目録,  
1604-1875



史料群記号: 33W 37S (5)

史料館  
国文学研究資料館

連絡先情報  
史料館  
国文学研究資料館  
東京都品川区豊町 1-16-10  
郵便番号 142-8585  
Tel: (03)3785-7131 (FAX)  
Fax: (03)3785-4436  
mailto:info@nlc.ac.jp  
http://nlc.nsl.ac.jp/

編集  
五島敏芳 Date completed  
Line 2000 EADコード化:  
五島敏芳, 2001

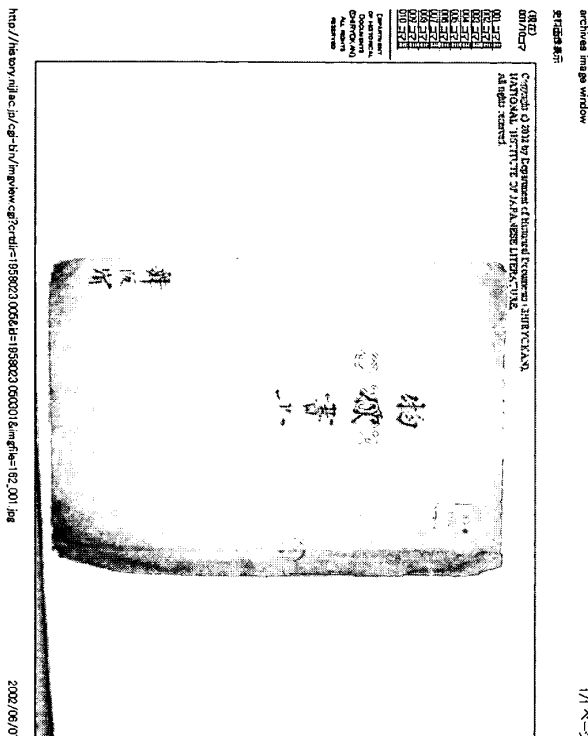
© 2006, 国文学研究資料館史料館

まえがき

このファイルには、史料館所蔵「遠江国敷知郡舞坂宿文書」の記録史料記述を明めた。「遠江国敷知郡舞坂宿文書」は、独立した史料群ながら、史料館所蔵「伊藤為之助収業文書」というコレクション史料に含まれており、コレクション史料の史料群を構成する1小群でもある。このてんを考慮し、「遠江国敷知郡舞坂宿文書」の史料群記号は、「伊藤為之助収業文書」の史料群記号に番号を加えて付与した。

x

史料館では、収業史料群を整理した成果を目録として刊行しているが、一部の史料群は規模の大小・整理計画・編集・予章等の都合から適切な刊行目録に成果が収録できないことがある。史料館では、目録未刊行で整理または仮整理(簡



(2). 山城国京都徳大寺家文書 (部分)

山城国京都徳大寺家文書目録, 1634-1932

5/105 ページ

奥則とその家族の原籍は神田区錦町1丁目2番地であった。同19年からは赤坂御所内真御門内官舎に奥則と三男則唐・三女藤子・四女治子・真母竹島・安八重の5人が移った。さらに、同24年8月に赤坂門内に官邸(閑院官旧邸)を与えられ、同月30日に神田本邸から赤坂官邸に移った(日記No.88)。奥則が官邸に居住する間に、神田邸は借地化が進んだ。

21は、明治26年12月29日に鷹司庶途から購入したもので、公弘の本邸となった。住所は後に、東京都下豊多摩郡千駄ヶ谷町488となった。この屋敷は史料上で「別邸」と呼ばれることもある。大正8年(1919)には、奥則が千駄ヶ谷邸において没しているため、奥則が内大臣を辞任した後は千駄ヶ谷邸に移ったものと見られる。

31は、天保10年(1839)に奥則が田中村の土地を買得たもので、徳大寺家では公統が居住した由緒ある土地と理解していた。明治40年(1907)に奥則の奥弟住友吉左衛門に譲渡された。これ以外には、茅ヶ畑に別荘がある。昭和初期には、渋谷区若木町15番地にも邸宅地があり、奥則の本邸とした。香堤寺は、十念寺(京都上京区寺町今出川)。奥則以降は東京谷川墓地に葬られ、後に多摩墓地に改葬された。

徳大寺家年譜

奥則  
天保10.12.6 誕生  
嘉永元. 7.12 教正五位下  
嘉永2. 1. 5 教正五位上  
嘉永3. 2. 3 教正五位下  
嘉永4. 3. 22 任侍従  
嘉永4. 7. 28 教正四位下  
嘉永4. 12. 4 元服(紫系色具段)  
嘉永5. 1. 27 教正四位上  
嘉永6. 5. 8 教正四位下  
嘉永7. 3. 7 任右近衛権少将  
安永4. 12. 8 任左近衛権中将有  
安永4. 12. 19 教正三位 中納言  
安永5. 3. 24 教正三位  
文久2. 2. 22 神皇正統記(使参向)  
文久2. 2. 27 復参  
文久2. 3. 11 賀茂行奉供事  
文久2. 4. 25 任権中納言  
文久2. 5. 21 隠居  
文久2. 5. 22 隠居衣

山城国京都徳大寺家文書目録, 1634-1932

20/105 ページ

直経院加印; 冊子(14.1×9.8) 1冊 32G 33T/1-5044  
印刷: [徳大寺]加印あり  
成吉思記上巻: 明治4年(1871)月2日; 6半 1冊 32G 33T/1-5011  
富岡政信, 明治廿三年十月廿九日出版

1-6. 書目録

32G 33T/1-6  
subref1

新題と内容

明治期以降に作られたと推定される書目録が2冊ある。なお、家持が作成した延喜書目録の中に、書目録があるため、3-3-7. 延喜書目録の項目を参照のこと。

清風館書目録; 半 1冊 32G 33T/1-6145  
録事あり  
【書目録】; 半野 1冊 32G 33T/1-6146

1-7. 有職故実

32G 33T/1-7  
subref1

新題と内容

いずれも、江戸期に作られた有職故実書である。「因幡八景 抄全」(No.407)の料紙は、京都市中の田代地所(現在の京都市中京区)のものを基に文書に用いており、奥則以下立派な家門で、一高連(千代八入三丁目)の地所がある。公統が作成した本家系図の文書が記述によって伝わり、「大空会感生主筆(即ち)奥則(神皇正統記)(No.306-3)は詳細な彩色絵図で、大空会感生主筆の故実、公統以下の記述などが記述され、挿絵で奥則会との親交が示されている。なお、公統の御印に、1-3. 記述の項を参照のこと。

因幡八景抄 全; 鹿野昌幸等; 6頁(2冊) 1冊 32G 33T/1-71402  
徳川家時傳 貞享1年(1684)八月十三日(徳川家時傳)御書(徳川家時傳)本  
享保十七日(1736)功之。信春(信春)之曰此本不事多之(信春)可(信春)也  
徳川家時傳; 1頁(1冊) 1冊 32G 33T/1-71398  
新巻祭: 寛政12年(1800)月2日; 6頁(2冊) 1冊 32G 33T/1-70218  
行幸召喚次第在中  
大空会三社奉拜(寛政)下上(信春) 共二(天明7年)甘藷寺(信春)長記  
文政元年(1818)小倉(信春)記; 6頁(2冊) 1冊  
[徳大寺文庫]「公統之由」印あり  
徳川家時傳(信春)大空会(信春)御書(信春)之信(信春)下(信春)可(信春)通(信春)公(信春)司(信春)  
次(信春)明: 嘉永元年(1824)月2日; 6頁(2冊) 1冊 32G 33T/1-7006-1

図5: 記録史料記述のEAD化のためのツール例

(3). EAD/XML 記録史料記述ファイルのソース ((2) の部分)

```

<archdesc type='inventory' level='fonds' langmaterial='jpn'>
  <did>
    <head> 【記録史料の概観】
    </head>
    <repository label='収蔵:'> 国立史料館
    </repository>
    <origination label='出所・作成:'>
      <corpname>徳大寺家</corpname>
    </origination>
    <unitid label=' 史料群記号:' countrycode='JP' repositorycode='DHDNIJL' id='
1957007
    >
      32G 33T
    </unitid>
    <unittitle label='タイトル:'> 山城国京都徳大寺家文書
    <unitdate label='年代:' type='inclusive' normal='1634-1932'>
      寛永 11 年～昭和 7 年
    </unitdate>
    <unittitle>
      abstract label='要約:'
      一部には江戸期の記録文書も含む、明治から昭和にかけての
      私的な記録文書を中心とした徳大寺家の文書
    </abstract>
    <physdesc label='物的状態:'> (数量) --
    <extent type="点数" unit="点"> 1,279
    </extent><br/> (形態の特徴)
      実則は、「実則」の丸印(直径1.4cm)、及び徳の異体字「憲」の丸印… (中略) …
    </physdesc>
    <physloc label='配架:'>
  </did>
  <bioghist>
    <head> 履歴
    </head>
    <p> 徳大寺家は、藤原北家閑院宮流の堂上公家で、藤原公実の5男実能(1… (中略) …
    </p>
  <chronlist>
    <head> 徳大寺家年譜
    </head>
    <caption> 実則
    </caption>
    <chronitem><date> 天保10.12. 6 </date><event> 誕生 </event></chronitem>
    <chronitem><date> 嘉永元. 7.12 </date><event> 叙従五位下 </event></chronitem>
    <chronitem><date> 嘉永 2. 1. 5 </date><event> 叙従五位上 </event></chronitem>
    (中略)
  </chronlist>
  <note> <出典> 明治十一年十二月廿五日太政官賞勲局江御差出履歴書控(207号, 208号), 「諸願届」
    (197号より作成.
  </note>
  <note>
    <bioghist>
      <scopecontent>
        <head> 範囲と内容
        <p> 本史料群は、年代的には明治から昭和にかけての記録文書が大部分を… (中略)…
        </p>
      </scopecontent>
      <arrangement>
        <head> 整理
        </head>
        <p> 1. 当史料群はこれまで、史料仮目録B, およびカード目録により、閲… (中略) …
        </p>
      </arrangement>
    </note>
  </scopecontent>
  (後略)
  
```

※「」は、そこにタブ文字があることを示す。

(図 5 の続き)

タの出現順（ファイルへの記述順）が保たれるように表示（ブラウザ等への出力表現）を工夫すれば、データの意味も体裁も失われなはずである。

以上のようにして、日本の記録史料記述の EAD/XML 化を進めた。もとより ISAD(G) 準拠の記述データがあり、ISAD(G) - EAD の対応表も存在していたことから、主要タグへの理解を進めた段階で、EAD アプリケーション・ガイド付録の例や既存の実例を参考にして作業を進めることは、容易ではなかったが比較的順調に進めることができた<sup>[17]</sup>。

この適用実験の結果でできあがった EAD 記録史料記述は、いくつかの改変が加えられていて厳密に EAD のデータとして正確とはいえないが、文字コードの問題を捨象すれば、一定の国際的な情報交換にたえるはずである。

かかる EAD 記録史料記述の効率的な発生については順調ではなく、あくまで一定の考え方と手順だけを提示できただけで、具体的現実的な問題解決の手段を開発できたわけではない。その考え方と手順は次項にゆずり、ここでは、単に EAD を適用することで生じた EAD 記録史料記述と、既存の電子データから変換することで生じた EAD 記録史料記述との、結果的な相違を簡単にながめておく。

図 5 の (1) と (2) は、同じスタイルシートを使っているため基本的な表示は同じだが、そのソースは微妙に異なっている。図 5 の (3) は、(2) のソースの一部で、(1) のそれよりも“雑”にできている。たとえば、属性値をくくるのにシングル・クォーテーションを使うのは不適切な用法かもしれない。またたとえば、タグにくくられたデータ内容と、タグとの間に、不自然な空白が見受けられる。属性値のクォーテーションと、属性値との間にも、同じく不自然な空白がある。これらの空白は、タブ文字である。いまのところブラウザで表示する際には問題は起こっていない。そのタブ文字は、EAD 利用上無害といえるが、使用の効果もある。(3) では各タグの階層的位位置（深さ）を視覚的に示すインデント（字下げ）に

も使われている。(1) のソースではそういった配慮を施していない。

これらの特徴は、あるアプリケーション・ソフトの利用に由来している。市販表計算ソフト（筆者の場合は Microsoft® Excel 97、以下「Excel」）のシートを利用しているためである。記録史料記述の電子データが既存する場合、一定の整形がなされている場合が多い。そういったデータの処理に対し表計算ソフトは、テキスト整形の諸ツールに比べ——少なくとも、そういったツール類の操作に不慣れた者にとって——操作が容易である。たとえば、記述本文中にはほとんど使用されないような文字をデリミタにして区切られている 1 行を 1 レコードとするデータを、デリミタごとにフィールドを分割することは、表計算ソフトやデータベース管理ソフトにとって難しいことではない。このような操作の容易性は、必ずしもコンピュータの操作に習熟しているとは限らない日本の記録史料管理主体にとって重要である。この点から、翻って既存の電子データの変換に限らず、記録史料記述を生成するための編成過程にあって、ISAD(G) 準拠の、または EAD 化に直結する、基礎データの生成にも応用できる、と筆者には思えたのである。

## 5. EAD/XML 化の考え方と手順

前述のとおり日本の記録史料記述への EAD/XML 適用実験からは、EAD/XML 化が ISAD(G) 準拠記述と同様に十分可能であることはまず確認できた。問題は、その効率的な生成に残っているが、コンピュータに関する専門的な知識を有さない者でも扱うことのできる一般的なツールによって多少は効率的な生成が可能になるという見通しは得た。その具体的な考え方・手順を再現する。

記録史料記述の EAD/XML 化ツールは、先に 3 で少しふれたが、日本語環境で動作はしても操作できるものは存在しない。英語に堪

|                                                                                                                               | A         | B                    | C                                                                                                                         | D                  | E   |
|-------------------------------------------------------------------------------------------------------------------------------|-----------|----------------------|---------------------------------------------------------------------------------------------------------------------------|--------------------|-----|
| 1                                                                                                                             | <archdesc | type='inventory'     | level='fonds'                                                                                                             | langmaterial='jpn' |     |
| 2                                                                                                                             |           |                      |                                                                                                                           |                    |     |
| 3                                                                                                                             |           | <did>                |                                                                                                                           |                    |     |
| 4                                                                                                                             |           | <head>               | 【記録史料の概観】                                                                                                                 |                    |     |
| 5                                                                                                                             |           | </head>              |                                                                                                                           |                    |     |
| 6                                                                                                                             |           | <repository label='  | 国立史料館                                                                                                                     |                    |     |
| 7                                                                                                                             |           | </repository>        |                                                                                                                           |                    |     |
| 8                                                                                                                             |           | <origination label=' | 出所・作成:'>                                                                                                                  |                    |     |
| 9                                                                                                                             |           |                      | <corpname>徳大寺家</corpname>                                                                                                 |                    |     |
| 10                                                                                                                            |           | </origination>       |                                                                                                                           |                    |     |
| 11                                                                                                                            |           | <unitid label='      | 史料                                                                                                                        | 1957007            | '>  |
| 12                                                                                                                            |           |                      | 32G 33T                                                                                                                   |                    |     |
| 13                                                                                                                            |           | </unitid>            |                                                                                                                           |                    |     |
| 14                                                                                                                            |           | <unittitle label='   | 山城国京都徳大寺家文書                                                                                                               |                    |     |
| 15                                                                                                                            |           | <unitdate label='    | 年 寛永11年～昭和7年                                                                                                              |                    | '>  |
| 16                                                                                                                            |           |                      | 1634-1932                                                                                                                 |                    |     |
| 17                                                                                                                            |           | </unitdate>          |                                                                                                                           |                    |     |
| 18                                                                                                                            |           | </unittitle>         |                                                                                                                           |                    |     |
| 19                                                                                                                            |           | <abstract label='    | 要 一部は江戸期の記録文書も含む、明治から昭和にかけての私的な記録文書を中心とした徳大寺家の文書                                                                          |                    |     |
| 20                                                                                                                            |           | </abstract>          |                                                                                                                           |                    |     |
| 21                                                                                                                            |           | <physdesc label='    | 物 (数量)                                                                                                                    |                    | --> |
| 22                                                                                                                            |           | <extent type='       | 点数                                                                                                                        | 1,279              |     |
| 23                                                                                                                            |           | </extent><br/>       | 【形態の特徴】<br>実際は、「実則」の丸印(直径1.4cm)、及び徳の具体字「徳」の丸印(直径1.2cm)を用いている。公統の印については、1-5、記録の項目を参照のこと。<br/> 銅収書綴、葉書・封書が多い。徳大寺家専用の郵便は半紙判の |                    |     |
| <eadheader /><frontmatter /><work-itemlist /><archdesc (fonds) /><bioghist-chronlist /><dsc-c(-series1) /><dsc-c(-series1) /> |           |                      |                                                                                                                           |                    |     |

図 6: 記録史料記述のためのワークシート (部分)

能でテキストエディタや諸テキスト処理ツールを自在に使いこなせる者はともかく、日常業務に追われる多くの日本の記録史料管理組織のなかで記録史料記述の EAD/XML 化を進めることには大きな困難をとまなう。そのため、市販表計算ソフトのような比較的一般的なソフトを用いた支援ツールが求められる。ただし、そこには一定の条件が必要だと考える。それは、記録史料管理組織が提供しようとする記録史料記述 = 記録史料目録の記述要素を明確にすることである [18]。より具体的には、サブ・フォンドからサブ・n - シリーズにいたるまでの中間的なレベルの記述について、必要最低限の記述要素を確定しておく、ということが現実的に重要な条件となる。

この条件をふまえた提案は、次のようになる。(1) 大きく分けて、フォンド・レベル、サブ・フォンド～サブ・n - シリーズ・レベル、アイテム・レベルそれぞれにおいて、記録史料記述表現に使用する限定された記述要素のためのデータ入力枠を設定する。(2) 各レベルごとにデータベースを作成する。(3) 各記述要素に対応する適切・適当な EAD タグを付与する。(4) EAD ファイルとして必要な情報を、テキストファイル上に総合する。

(1)～(4) の具体的な手順として、Excel とワードプロセッサ (以下「ワープロ」、筆者の場合 Microsoft® Word 97) か、テキストエディタを使用する場合を考えた。それぞれ次のようになる。

(1) あらかじめ、(a) EAD メタデータ部分<eadheader>、(b) 前付け部分<frontmatter>、(c) フォンド記述部分<archdesc>、(d) 付属資料 (年譜など) 部分<chronlist>など、(e) 中間的記述 (series など) 部分<dsc><cXX>、(f) アイテム記述部分、のそれぞれのワークシートを Excel のブックに用意し、テンプレートとする。(c)・(d)・(e)・(f) については、各記述レベルにそくした記述要素を設定し、入力部を明示 (色分け) し、タグを前後に配置——シートのデータ入力面をコピーしたときに、入力した記録史料記述データとタグが、デリミタとともに一括されるように——する (図 6 参照) [19]。

(2) 既存の電子データがある場合、当該電子データのファイルから、Excel の適切な記述部分シートへ、データを複写・貼付する (図 7 参照)。整理による記録史料の基礎データ採取の過程で発生するデータを編集する場合、アイテム・レベルのシートにあらかじめ設定した



| A | B        |                                            |  |  | C           | D          | E       |
|---|----------|--------------------------------------------|--|--|-------------|------------|---------|
| 1 | Org-data |                                            |  |  | Div-element |            |         |
| 2 | Sg       | DSCP                                       |  |  | unititle    | contro     | corpnam |
| 5 | 3        | 「徳大寺殿御由緒書」(No. 237)は、文化6年(1809)段階での23代実祖(1 |  |  |             | #VALUE!    | #VALUE! |
| 6 | 4        | 徳大寺殿御由緒書(実祖～実堅)★文化8年★美 1冊★237              |  |  |             | 徳大寺殿御由緒書(実 | #VALUE! |
| 7 | 5        | 御家記★徳大寺家、徳大寺家10行割紙 表表紙裏「悪」印あり★明治7年1        |  |  |             | 御家記        | 徳大寺家    |
| 8 | 6        |                                            |  |  |             | #VALUE!    | #VALUE! |

| F       | G       | H       | I       | J       | K      | L      | M         | N       | O       | P       | Q              |
|---------|---------|---------|---------|---------|--------|--------|-----------|---------|---------|---------|----------------|
| corpnam | subjec  | unitda  | physdes | unitid  | [paren | [branc | [refcode  | [@id-fo | [@id-sr | [@id-it | [@id]          |
| #VALUE! | #VALUE! | #VALUE! | #VALUE! | #VALUE! |        |        | \$2G \$3T | 1957007 | 001001  |         | 1957007.001001 |
|         | 文化8年    | 美       | 1冊      | 237     | 237    |        | \$2G \$3T | 1957007 | 001001  |         | 1957007.001001 |
|         | 徳大寺家、   | 明治7年1   | 半野      | 1冊      | 22     | 22     | \$2G \$3T | 1957007 | 001001  |         | 1957007.001001 |

| R             | S                                                                                               | T       | U       | V       |
|---------------|-------------------------------------------------------------------------------------------------|---------|---------|---------|
| for-copy-data |                                                                                                 |         |         |         |
| [level]       |                                                                                                 |         |         |         |
| 1-1.          | #VALUE!                                                                                         | #VALUE! | #VALUE! | #VALUE! |
| 1-1.          | < level='item'><did:&unititle id='1957007.0010010004'>徳大寺<physdes>美 1冊</&controlaccess>&corpnam |         |         |         |
| 1-1.          | < level='item'><did:&unititle id='1957007.0010010005'>御家記<physdes>半野 1冊<controlaccess>&corpnam  |         |         |         |

※実際にデータを入力しなければならないセルは、色付けされた特定のセルのみ。

図7: アイテム・レベル記述のためのワークシート (部分)

記述要素の枠に沿ってデータを蓄積する——適宜、作業ノートとして記録しておくような情報は、最終的な EAD 表示に矛盾しない位置を想定し<note>タグの間に入力する。後の記録史料記述表示のための作業効率化のため、基礎データ採取の処理段階から、EAD 記述要素を意識すべきである。このように Excel のブックに記録史料記述データベースを作成する [20]。

(3) Excel のブックにおいて記録史料記述が完成したら、テキストファイルへ各シートのデータを複写・貼付する。各記述要素のデータの前後に、当該記述要素に適切なタグが位置するよう、注意しながら作業する。ファイル容量を節約する場合、Excel からのコピー時にデリミタとしてタブ文字が挿入されるため、これをワープロやテキストエディタを用いて、一括削除(空文字との置換)をおこなう。アイテム・レベルの記述情報は、Excel のテンプレートでは、すべて<c>タグを使用しているため、適切な階層の小群記述に配置する際に、当該小群記述の1レベル下位の数値を付与する形に編集する [21]。

(4) EAD/XML ファイルとしての体裁を整え、当面の表示に使用するスタイルシートの記述等のいくつかのメタデータを付与する。

前項③(2)の事例は、上述のような過程を経て EAD/XML となったものである。このよ

うな EAD/XML 記録史料記述を基本とすれば、そこから記録史料へ接近するための多様な検索システムを構築できよう。たとえば、各記録史料群の EAD/XML ファイルから、XSLT を使ってフォンド・レベル記述のみを抽出し、それらを統合することで、当該記録史料群を収蔵する記録史料管理組織の収蔵記録史料ガイドを作成することができる。

記録史料管理のための基本的な情報ともなるはずである。たとえば、同様 XSLT によってアイテム・レベル記述のみを抽出し、請求記号(配架番号)順に並べ替えれば、配架順目録を生成することができよう。

このような効果を考えたとき、あらためて問題となるのは、記録史料記述の情報の維持のあり方である。たとえば、修正しなければならないような記述内容上の誤りがあった場合の編集作業は、どの電子データに対しておこなえばよいか、といったことが直接の疑問として想定できる。

記録史料記述の情報の維持は、対象となる記録史料の状態によって変化しよう(図8参照)。すなわち、それが記録史料記述を発生する過程にある＝編成途中にあるのか、それとも記録史料記述が一応の完成をみている＝編成停止にあるのか、によって EAD/XML 化の予定がある記録史料記述のデータ保持の段階が異なるのである。

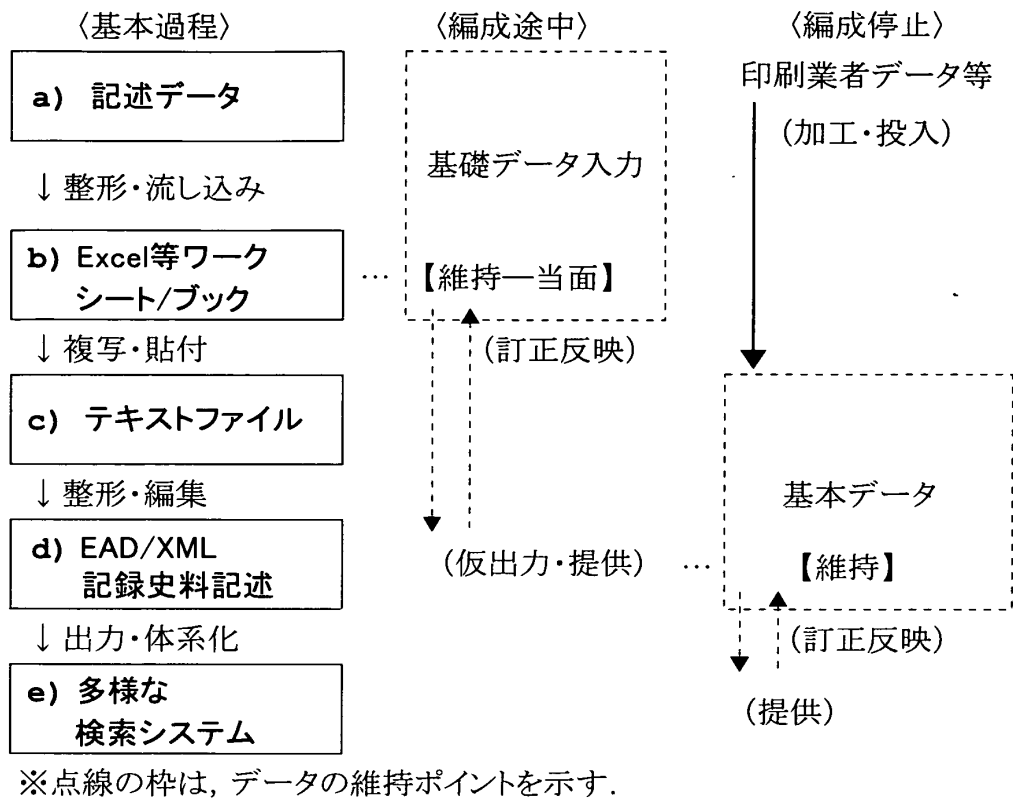


図 8: EAD/XML 記録史料記述の生成と維持

いま示した Excel のワークシートや他の同等の機能を持つデータベース管理ソフトを使用して EAD/XML 記録史料記述を生成するものとして、まず当該の記録史料が編成停止状態にある場合——たとえば、既に刊行目録のような（一応）完成した記録史料記述がある場合、さらにその場合で刊行目録の電子データが刊行を担当した印刷業者から入手できる場合——は、基本データは EAD/XML 化した記録史料記述におくべきで、その修正編集は EAD/XML のテキストファイルに対して直接操作する方がよい。テキストファイルであるがゆえに、特定のアプリケーション・ソフトに依存せず、保存性が高いからである。

次に、編成途中にある場合、しばしば記録史料記述は完成しているわけではない。しかし、ひと通りの物的な措置——保存措置がなされ、閲覧に供することのできる状態にあれば、仮の検索手段を提供することがあるかもしれない。記録史料を閲覧提供した結果、あたらしい情報をもたらされ、当該記録史料の

編成が進展したり、仮編成の状態が大幅に変更されることもあるだろう。その場合、編成作業で使用している Excel なりデータベース管理ソフトにおけるデータ＝すぐに編成状態を変更することが可能なデータの形で保持しておくことが有効である。

このように考えると、記録史料管理のなかで編成過程にある記録史料の記述の情報は、なるべく将来の EAD/XML 化における再利用を想定・意識して、EAD または ISAD(G) の情報要素と同じ内容であった方がよい。記録史料記述（とくに記録史料目録）の表現において限定された各記述レベルごとの記述要素であれば、EAD または ISAD(G) の情報要素は、記録史料の編成に際して採取される基礎データとしても十分に利用可能ではないだろうか。日本における記録史料編成の現実的な作業のうち、とくに記録史料の持つ内的秩序の復原については、もっぱらアイテム・レベルの基礎データ採取が主要な位置を占める。アイテム・レベルにおける EAD の、

少なくとも表現上の必要から限定された情報要素は、基礎データ採取の要素として従来から用いられてきた要素と適合的ではないかと考えている。

日本の記録史料管理組織では、記録史料編成支援上の表計算またはデータベース管理ソフトの利用に蓄積がある。そういった表計算またはデータベース管理ソフトの利用に、EADの情報要素を組み込むことを提案したい。効率的なISAD(G)準拠の記述またはEAD化された記述の作成は、コンピュータ上の検索手段を作り直すような手間を省き、記録史料管理の過程全体のなかで記述に割く比重を軽くし、よい作業バランスをもたらすはずである。

## 6. おわりに

日本の記録史料記述に対するEAD/XMLの適用実験は、単にその適用可能性を証明するものではない。あたらしい技術の導入が、記録史料管理の過程を改変する可能性をも示している。これまで用途によって分断されていた記録史料管理に関する諸情報を、ひとつに結びつける役割を担うことができる手段として、EAD/XMLを理解することができるのではないかと考えている。

なお、記録史料記述のEAD/XML化を日本で普及させるには、あまりに現在のデータ生成手段は貧弱であるし、その即物的な効果が情報の国際交換の可能性と電子化された記録史料記述の保存への有用性にしか明確にできていない。また記録史料の利用者の利便性を確保するはずの、EAD/XMLをもとにした多様な検索システムの構築については、筆者の不勉強と技術的な知識の限界からモデル・具体像とも提示できなかつた<sup>[22]</sup>。自ら今後の課題としつつも、活発な議論とご教示を乞う次第である。

## 謝辞

本稿の作成にあたり、まず国文学研究資料

館史料館と館員諸氏のご理解に感謝いたします。そして、EAD利用の有効性を説得していただき数々の資料と助言を提供して下さった安澤秀一先生、成稿までに叱咤激励して下さった安永尚志先生、ほかいろいろとご教示ご指導くださった諸先生・諸兄弟の皆様に感謝申し上げます。ありがとうございました。

## 参考文献

- [1] アーカイブズ・インフォメーション研究会編訳『記録史料記述の国際標準』（北海道大学図書刊行会、2001年2月）での訳語。もと、General International Standard Archival Description: ISAD(G)。ごく簡単に説明を補足すれば、次のとおり。国際文書館評議会 International Council on Archives: ICAにより定められ、1994年には初版が公刊された。その内容は5年ごとに見直されるものとされ、2000年9月に第2版が公表されている。
- [2] 筆者が把握しているものでは、次の諸成果をあげることができる。青山英幸「〈報告〉国際標準記録史料記述等による箱館奉行文書目録作成の実験について」『北海道立文書館研究紀要』第12号、1997年。森本祥子「国際標準記録史料記述（一般原則）適用の試み——諸家文書の場合」『史料館研究紀要』第28号（国文学研究資料館史料館）、1997年。森本祥子「国際標準記録史料記述（一般原則）適用の試み——行政文書の場合」『史料館研究紀要』第29号（国文学研究資料館史料館）、1998年。安藤正人『記録史料学と現代——アーカイブズの科学をめざして』吉川弘文館、1998年。佐々木和子「第20航空軍・第21爆撃機軍団戦記史料の目録記述の試み——記録史料学からみたピースおおさか史料」『戦争と平和'01』第10号（大阪国際平和研究所紀要）、2001年。
- [3] 論文として公表された成果として、小暮隆志・鈴木一哉「群馬県立文書館ホーム

ページ上での館所蔵文書紹介の試み」(群馬県立文書館『双文』第18号, 2001年)をあげる。ただ, HTML文書での適用にとどまり, 電子データとしての構造化は考慮されていない。なお, 既にホームページでISAD(G)に準拠し, あるいはそれを意識したデータベースを公開している実践としては, さしあたり次の各機関の例をあげておく。

国立公文書館アジア歴史資料センター  
(<http://www.jacar.go.jp/>)

一橋大学附属図書館・デジタルアーカイブ  
([http://www.lib.hit-u.ac.jp/service/index\\_Ja.html](http://www.lib.hit-u.ac.jp/service/index_Ja.html))

国文学研究資料館史料館・史料情報共有化データベース  
([http://history.nijl.ac.jp/db/akyoyu/akyoyu\\_top.htm](http://history.nijl.ac.jp/db/akyoyu/akyoyu_top.htm))

なお, 本稿で参照を指示しているホームページ等のURLは, いずれも2002年8月現在に確認したものである。

- [4] 五島「XMLを利用した史料記述の可能性——『国際標準：記録史料記述の一般原則』ISAD(G)第2版とデータベースをめぐって」情報知識学会人文社会科学系部会「歴史研究と電算機利用ワークショップ」口頭報告, 2001年7月14日(東京)。
- [5] SAAのEAD作業部会と, アメリカ議会図書館の関係事務局が, 維持にあたっており, とくに後者はホームページ等を通じてEAD関係ファイルの配布している(<http://www.loc.gov/ead/>)。もと, 1993年開始のカリフォルニア大学バークレイ校図書館での機械可読検索手段のための符号化標準規格開発の研究プロジェクトが前身と思われる。
- [6] たとえば, 次の2報告をあげる。青山英幸「国際標準とEAD」アーカイブズ・インフォメーション研究会主催「オープンセミナー・インきょうと'2000」口頭報告, 2000年12月3日(京都)。安澤秀一「エンコードアーカイバルデスクリプションEAD:SGML/XMLの応用形として」情報処理学会『人文科学とコ

ンピュータ』51-3, 2001年7月13日(埼玉)。とくに安澤氏は, 上掲報告以外にも, いくつもの関連報告があり, 精力的にEADを紹介している。

- [7] 現在, 国文学研究資料館史料館のホームページにおいて, 一部の成果を公開している。次のURLを参照のこと:  
<http://history.nijl.ac.jp/dhd2.5.htm>
- [8] EADのアプリケーション・ガイドラインの付録Bに「EAD Crosswalks」に対応表が示されている。ISAD(G)との対応について付言すると, 次の点に注意しなければならない。ほぼすべてのISAD(G)の記述要素がEADのタグと対応しているが, EADによる記録史料記述での最低限必要なタグ(記述要素)がISAD(G)の記述要素すべてを含んでいるわけではない, という点である。なお, 上記アプリケーション・ガイドラインや, 本章での議論の素材としたEADのDTDファイル類は, いずれも既掲注[5]のアメリカ議会図書館の維持するホームページから入手した。
- [9] 一般的に記録史料は群として存在し, 物的な1点にいたるまでの間に, しばしばいくつもの小群でまとまっている。この在り方を把握し記述する際, たとえば特定の団体の記録史料群全体に対応する記述レベルは「フォンドfonds」であり, 記録史料の物的な1点「アイテムitem」(管理上の最小単位)となり, しばしばその記録史料群を形成した当該団体の内部構成に対応して, 小群の記述レベル「サブ(^n) - フォンド sub(^n)-fonds > シリーズ series > サブ(^n) - シリーズ sub(^n)-series > ファイル file」を設定できる(不等号は小群の記述レベルの大小の関係を示す)。各小群をあらわすサブ(^n) - フォンドからファイルまでは, 物的な単位と対応することもあれば理念的に復元された単位であることもあり, 途中のレベルが欠けることもあり「サブ-」「サブ-サブ-」…と複雑に増える

こともある。

- [10] 筆者の所持するパーソナルコンピュータの環境で、実験に利用したブラウザは、Microsoft Internet Explorer (以下 IE と略) version 6 であった。他のブラウザとして、Netscape 6 もあったが、ごく初歩的な XML 文書の表示実験において、CSS を通した場合は表示できたが、XSL を使用した場合は表示できなかったため、同ブラウザを使用しなかった。
- [11] 情報要素 (記述項目) を羅列的に表示するに近い、データ確認のためだけに使用できるような CSS は存在する。これをもとに体裁を整えることは、重要な作業かもしれない。
- [12] 次の URL を参照のこと：  
<http://cidc.library.cornell.edu/xml/>  
 なお、当該のページの冒頭に表示される組織名は、「Cornell Institute for Digital Collections」となっている。
- [13] 次の URL を参照のこと：  
[http://www.chs.org/library/ead\\_proj.htm](http://www.chs.org/library/ead_proj.htm)
- [14] 脈絡を無視して利用すれば、利用された記録史料は、当該の利用によって期待された結果や効果の根拠・証拠としての位置を失うことになろう。かかる根拠・証拠は、記録史料目録などの検索手段ないし記録史料管理主体そのものによって保証されていることを忘れてはならない。
- [15] たとえば、Jackie M. Dooley, ed., *Encoded Archival Description: Context, Theory, and Case Studies*, Society of American Archivists (Chicago), 1998 所収論文では、Author/Editor や WordPerfect's SGML editor 等の語を目にすることができ、それらが活用されていることを知る。そこで示されている EAD ファイルやそのテンプレートの事例からも、タグの配置や用法が少しずつ異なっていることを推察できる。
- [16] イギリスのマンチェスター大学がホスト役となって、現在いくつかのイギリスの大学間で動いていると思われる。次の URL を参照のこと：<http://www.archiveshub.ac.uk/>
- [17] このとき、前項にふれたアーカイブズ・ハブ (注 [16] 参照) が提供する、EAD ファイル生成フォーム (Javascript を使用しているらしい) も、実際に試用した。EAD に詳しくなくてもブラウザ画面の操作から生成でき、日本語文字列も入力可能であった。フォンドおよびサブ・フォンド・レベル記述の生成には有効である。
- [18] たとえば、フォンド・レベルからアイテム・レベルにいたるまで、ISAD(G) の全記述要素で記述し提示することは可能ながら、その表現の体裁が全ての記述レベルにわたって ISAD(G) の全記述要素を提示するものでよいかどうか、ということである。
- [19] Excel のセルは、クリップボードへコピー可能な入力文字数が限られているため、テキストが大きければ入力データの前後をタグがはさむ形を維持しつつ、適宜複数のセルにまたがらせる。
- [20] 記録史料記述の詳細部分 <dsc> における、中間的記述に複数の階層がある場合、より包括的な記述から順に <c01>・<c02>・<c03>… とレベルを混同しないように入れ子にしていくように注意する。
- [21] たとえば、シリーズ・レベル記述が <c03> タグ内にあった場合、その直下のアイテム・レベル記述は、それぞれ <c04> タグ内に収められる。あらかじめ特定の階層にしかアイテム・レベルがない場合、Excel シート上でアイテム・レベルでの使用タグを <c> から <cXX> に変更しておくこともできる。  
 なお、アイテム・レベル記述の階層変更の一括処理例として、Microsoft® Word 97 での正規表現置換の場合をあげておけば、次のようになる (それぞれ検索文字列と置換文字列として入力する値を示す)：  
 「(¥<)c(level='item'¥>)(\*) (¥</)c(¥>)」 → 「¥1c03¥2¥3¥4c03¥5」

(<c~>タグ →<c03>タグへ置換).

また, mule など emacs 系のエディタなら, 適切な位置から M-x-replace-regexp によって正規表現置換を起動し, ミニ・バッファへ順に次のように入力する(タイプライタ字体の部分): 「Replace regexp: ~¥(<¥)c¥(.+¥)¥(</¥)c¥(>¥)」 → (検索文字列確定) → 「Replace regexp: ~¥(<¥)c¥(.+¥)¥(</¥)c¥(>¥)\$ with: ¥1c03¥2¥3c03¥4」(下線部分を任意のレベルへ変更する) → (置換実行).

- [22] 検索システムについては, 注 [3] の史料館・史料情報共有化データベースに関する公開研究会での報告(2002年1月10日・東京, 平成11~13年度科学研究費

補助金基盤研究B(2)「歴史史料情報の共同集約と共有化に向けてのシステム構築に関する研究」[研究代表者: 鈴江英一, 課題番号: 11410096] 研究成果報告書所収「『史料情報共有化データベース』における横断的検索実験」参照)で示した内容をもとに, ごくおおまかなイメージを「日本における記録史料記述 EAD/XML 化の実践」(情報処理学会人文科学とコンピュータシンポジウム・じんもんこん :-)2002, 2002年9月22日・大阪)で, あらためて提示した.

(2002年9月30日受付)

(2002年11月22日採録)

# バイオ研究者間の情報共有のためのプロテオーム XML 記述方式 An XML Format for Proteomics Database to Accelerate Collaboration among Bioresearchers

水口 弘紀<sup>†</sup> 上條 憲一<sup>‡</sup> 次田 皓<sup>‡,\*</sup>

Hironori MIZUGUCHI<sup>†</sup> Ken'ichi KAMIJO<sup>‡</sup> and Akira TSUGITA<sup>‡,\*</sup>

プロテオーム解析におけるデータ共有のための XML 記述方式を提案する。さらに、本 XML 記述を使ったプロトタイプを開発した。プロテオーム解析は、細胞内で発現しているタンパク質の網羅的同定やタンパク質の機能解明を目的とする。プロテオーム解析では、解析対象となる試料情報、タンパク質の分離手法、解析結果である分離されたタンパク質名称、機能情報などを扱う。特に、解析データを共有、交換し、実験の再現や比較を行うには、実験結果だけでなくタンパク質の分離手法などの実験手順情報も重要となる。我々は、バイオ研究者間のデータ共有をより簡単に行うことを目指し、プロテオーム解析結果だけでなく、実験手順情報も記述できる XML を用いたデータ記述方法を提案する。さらに、実験手順に応じて情報を入力し容易に XML データを作成できるエディタと、この XML データを登録検索できるデータベースサーバシステムを開発した。

We propose an XML (eXtensible Markup Language) format for proteomics database to exchange proteome analysis data. The XML-based data is highly machine-readable and easy to represent information hierarchy and relationships. There have been several XML formats of proteome data which mainly represent the sequence information stored in the Protein Identification Resource (PIR) and the Protein Data Base (PDB). Our XML-based data format has a proteome-analysis-oriented structure and describes information of sample preparation, 2D gel electrophoresis images, spot identification information in the gels and the sequence information of the spots. The model is used to exchange both of preparation parameters and the results of 2D gel electrophoresis analysis. It would accelerate collaboration among proteomics researchers if a platform exchanging these data is developed on the internet. By using the XML-based data format for proteomics, we have developed an XML editor and a web-based prototype system which consists of XML database, agent, security and graphical user interface (GUI).

キーワード：バイオインフォマティクス，プロテオーム解析，情報共有，XML，データベース  
Bioinformatics, Proteome, Information sharing, XML, Database

## 1. はじめに

近年、バイオテクノロジーやライフサイエンスの分野では、データベースを利用した情

報蓄積、インターネットを利用した情報共有を盛んに行っている。データベースのエントリー数も毎年増加しており、遺伝子の公開データベースの DDBJ, EMBL, GenBank などでは、年に 2 倍のペースで増加している<sup>[1]</sup>。バイオ研究者は、新たに発見した情報を論文発表する前に、関連する遺伝子やタンパク質の情報を公開データベースへ登録することが義務付けられている場合が多く、データベースの重要性は増す一方である。バイオ研究は、このような公開データベースを検索し、新しい

<sup>†</sup> NEC インターネットシステム研究所, Internet Systems Research Laboratories, NEC Corporation, 奈良県生駒市高山町 8916-47

E-mail: hironori@ab.jp.nec.com

<sup>‡</sup> NEC 基礎研究所プロテオミクス研究センター, Proteomics Research Center, Fundamental Research Laboratories, NEC Corporation, 茨城県つくば市御幸が丘 34 番地

E-mail: k-kamijo@az.jp.nec.com

\* E-mail: a-tsugita@ce.jp.nec.com

情報を取得し、取得した情報を利用して新たな発見が生まれる、というサイクルによって発展していると言っても過言ではない。

プロテオミクスは、細胞中で発現<sup>1</sup>しているタンパク質(プロテイン)の網羅的な同定(タンパク質名称のマイニングなど)や、タンパク質立体構造の同定、タンパク質の機能解明などタンパク質に関わる研究分野である。細胞や組織などの試料から、タンパク質を分離し、分離されたタンパク質の名前を同定し、機能を調べることをプロテオーム解析と呼ぶ。タンパク質は、20種類のアミノ酸を組み合わせた一次元配列で表されており、その配列情報やタンパク質の名前を知ることが、機能解明の一步となる。また、タンパク質は臓器や病気に応じて細胞内に存在する種類が異なっていることがわかっており、病気解明や創薬への利用が注目されている。タンパク質の種類は非常に多く、たとえば、ヒトの場合、約4万種類の遺伝子から作られるタンパク質は選択的スプライシングや修飾、切断などの生化学的な作用によって、10万種類以上存在すると言われている。研究者は、これらの遺伝子やタンパク質データを公開データベースに登録し、共有することで、調査対象の細胞に存在するタンパク質の同定、その立体構造や機能の解明につなげていく。

従来、バイオ分野で多く用いられてきたデータフォーマットは、特殊な例を除きテキストファイルであり、ほとんど構造を持っておらず、データベース毎に記述方法も異なっている。このようなデータは、一般のテキストエディタで編集することができ、研究者にとって見やすいという利点がある。しかし、コンピュータ処理に必要なデータ交換のためには、データベース毎に特別な読み込みなどの処理を用意しなければならず、不向きである。

我々は、プロテオーム解析の結果などを研究者間で共有し、プロテオミクスを促進させるためのXML記述方法を提案した。XMLは、

テキストファイルとしてデータを保存でき、しかも、構造化されているのでソフトウェアでの処理も容易である。このため、研究者が見て理解でき、コンピュータ同士のデータ交換にも適している。

提案するXML記述方法の特徴は、実験結果のタンパク質データだけでなく、実験過程(実験プロトコルとも呼ばれる)を記述できる点である。プロテオーム解析における実験は、実験の再現性が重要であり、実験に使われた試薬や、調合方法、温度などの実験プロトコルに関する情報を記述することにより、実験の信頼性を判断したり、追試を行ったりすることが可能となる。また、それらの情報がプロテオーム解析の結果に付加されることにより、解析結果の信頼性が向上したり、自らの解析結果との比較を行ったりすることができる。

さらに、提案するXML記述方法を容易に記述するエディタと、このXMLデータを登録、検索し共有するためのデータベースサーバシステムを開発した。エディタは、プロテオーム解析の実験手順毎に入力を行うことで簡単にデータを作成できる。データベースサーバシステムは、検索項目を細かく指定可能であり、実験結果だけでなく実験プロトコルに利用した試薬までも検索可能である。

## 2. プロテオーム解析のインフォーマティクスにおける課題

プロテオーム解析は、多くの公開データベースを利用して、タンパク質の同定や機能発見を行う。これら公開データベースのエントリは、一般にフラットファイルフォーマットと呼ばれる構造化されていないテキスト形式で公開されている。タンパク質データベースで世界的に有名なPIR<sup>[2]</sup>とSWISS-PROT<sup>[3]</sup>のエントリ例を図1に示す。左がPIR、右がSWISS-PROTのエントリであり、同じタンパク質データを表している。双方とも、各行がデータ項目をあらわしており、行頭に項目名、次に空白、次に項目内容を記述している。

<sup>1</sup> 発現とは、細胞内の遺伝子がタンパク質を作り出すことをいう。



PIRの内容は以下である。1行目に識別子, 2行目にタンパク質の機能概要を示す意味のわかるタイトル, 3行目から5行目に生物種名と生物種データベースへの参照, 6行目から8行目にデータベースエントリーの作成更新日時, 9行目にデータベースへアクセスするための番号, 10行目から論文などへの参照, 次にアミノ酸配列が示されている。SWISS-PROTの内容は以下である。1行目に識別子, 2行目にデータベースへアクセスするための番号, 3行目から5行目に作成更新日時, 6行目に機能概要, 7行目に関連遺伝子名, 8行目に生物種, 次に参照情報, アミノ酸配列情報が示されている。

この例から次の二つの課題が挙げられる。

- データフォーマットが異なる
- 実験過程情報の欠如

図1の二つのデータベースのデータフォーマットは, 各行がデータ項目を表すことや, 識別子, 論文への参照, アミノ酸配列を持つことなどの共通性はあるものの, 項目の数や順序も異なり, 項目の意味は同じであるが項目名が異なっているなど記述形式が異なる。このようにフォーマットが異なっていると, データ内容や項目の違いを理解せねばならず研究者間でのデータ交換を行って, コンピュータで管理することは難しいと思われる。研究者間で同じフォーマットのデータ, もしくは,

フォーマットが簡単に変換できるデータを共有する必要がある。

一般にデータベースのエントリー内容は, 実験対象となる生物種や組織名, 機能, 発表論文への参照など実験結果や解析結果の情報のみであり, 実験過程の情報は多種多様な手段を用いているため含まれていない。実験過程は, 実験結果や解析結果データの信頼性の判断や実験結果の比較のためには必要である。特に, プロテオーム解析では, 実験に利用する試薬は研究者ごとに異なる場合があり, 実験対象が同じであっても試薬の違いによって実験結果が異なる。したがって, どのような手順, 試薬を利用したか明確に記述することは, 実験結果の信頼性を向上させる上で必要である。

### 3. 課題解決へのアプローチ

これらの二つの課題に対し, 実験過程を記述できるプロテオーム XML 記述形式を提案する。XMLで記述することで, データ共有に利用しやすいフォーマットとなり, フラットファイルの課題を解決できる。XMLの特徴である, 人間も計算機も読み書き可能である点, データ交換の容易さ, XSLTを利用したデータ構造の簡単な変換ができることなどを有効に活用することができる。さらに, 提案

|                                                                                                                                                                                                                                                                                                                                |                                                                                                                                                                                                                                                                                           |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <pre> ENTRY      xxxxx TITLE      G protein-coupled xxxx ORGANISM   #formal_name Caenorhabditis elegans   #cross-references taxon:6239 DATE       29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change      29-Oct-1999 ACCESSIONS  xxxxx REFERENCE   Z21511   #authors      (省略) (省略) SEQUENCE (省略)                 </pre> | <pre> ID xxxxx AC xxxxx DT 01-NOV-1995 (Rel. 32, Created) DT 15-DEC-1996 (Rel. 33, Last sequence update) DT (省略) DE Serpentine receptors xxxxx GN SRG-9 OR T12A2.10. OS Caenorhabditis elegans. (省略) RN [1] RP SEQUENCE FROM N. A. RC STRAIN=(省略); RA (省略) SQ (省略)                 </pre> |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

図1: タンパク質データベース PIR (左) と SWISS-PROT (右) のエントリー例 (図中“(省略)”の部分は紙面の都合のため省略)

する記述方式には実験結果だけでなく、実験過程も記述できるため、実験の比較や妥当性の検証などを行うことができる。

最近、バイオ分野ではデータベース間でのフォーマットの違いを吸収するため、公開データベースのエントリをXMLで記述することが行われている。また、データ解析ソフトウェアでのXML利用も行われている。

遺伝子データベース GenBank<sup>[4]</sup>, EMBL<sup>[5]</sup>, DDBJ<sup>[6]</sup>, タンパク質データベース PIR<sup>[2]</sup>は、データベースエントリをXMLで提供するサービスを開始している。これらは、既存のデータエントリをXMLで記述しユーザに提供している。しかし、これらには、実験過程があまりに多種多様なため、実験を再現する実験過程情報は含まれていない。

データ解析ソフトウェアで利用するために独自に決めたXML記述方法として、BSML<sup>[7]</sup>, BIOML<sup>[8]</sup>がある。これらは、ソフトウェアで独自に決めたフォーマットであり、主に遺伝子塩基配列やタンパク質アミノ酸配列データを格納するために用いられる。データは配列データのみである。

実験過程の情報をXML化しようという試みは、プロテオミクス分野ではなく、遺伝子の発現解析分野で行われている。たとえば、データ解析ソフトウェアと実験機器で交換するためのXML記述方法としてMAGE-ML<sup>[9]</sup>

がある。これは、遺伝子解析用の実験機器であるマイクロアレイに用いられるフォーマットであり、実験結果を解析するためのソフトウェアとのデータ交換が目的である。MAGE-MLには、実験過程情報も含まれているが、遺伝子の機能などを解析するゲノム研究に適用するフォーマットであり、マイクロアレイ機器に特化したフォーマットである。

#### 4. プロテオーム解析の実験プロトコル

プロテオーム解析は、創薬や診断のマーカーのための標的タンパクを探したり、興味ある組織で発現しているタンパク質を網羅的に調べたりする目的で使われる。例えば、正常細胞で発現しているタンパク質と疾患細胞で発現しているタンパク質の集合を比較して、標的タンパク質を探したり、同じ組織でも、時間的経過によるタンパク質発現の変化を調べたりすることができる。

図2にプロテオーム解析の実験過程の一例を示す。一般に、プロテオーム解析は、タンパク質を分離するための二次元電気泳動法や液体クロマトグラフィ法、タンパク質を同定するためのペプチド・マス・フィンガープリント法、ペプチド・シーケンス・タグ法、N末端シーケンシング法、C末端シーケンシン

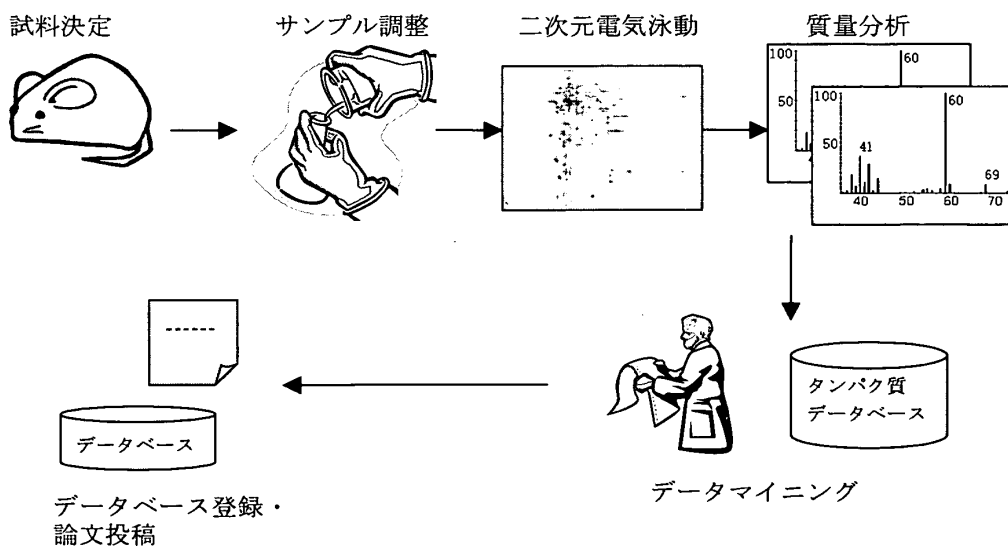


図2: タンパク質同定における実験過程

グ法，フラグメントペプチドのマス／マス分析で，そのペプチドの両末端のシーケンス法などを行う．以下では，二次元電気泳動法と，ペプチド・マス・フィンガープリント法を用いた実験過程を説明する．

研究者は，最初に，研究ターゲットとなる試料を決定する．試料とは，発現しているタンパク質を調べたい細胞や組織である．研究対象となるヒトやイネといった生物種，肝臓や腎臓，あるいは葉や根といった部位，病気や発芽といった状態から試料を決定する．

次に，試料からサンプル調整を行う．サンプル調整とは，試料となる細胞や組織からタンパク質を溶液中に抽出する工程である．細胞は細胞膜に覆われているため種々のデタージェントを用いたり超音波などを用いて破壊し，さらに，細胞にはタンパク質以外の物質も含まれているため様々な試薬や遠心分離機などを利用して分離抽出する．詳細は，参考文献<sup>[10][11][12]</sup>に紹介されている．試料に応じて，適切な細胞破壊の方法やタンパク質抽出の試薬などがある．

サンプル調整で得られたサンプルには多く

の種類タンパク質が溶け込んでいるため，次に，これらを二次元電気泳動法によりタンパク質を性質別に分離する．サンプルを電荷と分子量というタンパク質の性質別に分離するのが二次元電気泳動法であり，これにより3000種類以上のタンパク質を一度に分離できる．この実験は4つのステップに分かれる．1) サンプルを流し入れ棒状ゲルに接触させ，ゲルの両端に電極を付け，適切な電圧をかける．これにより，タンパク質は電荷に応じて対応する位置に移動する．2) 垂直に立てた平面状のゲルに，この棒状のゲルをのせ，上下に電圧をかける．これにより分子量の違いによってタンパク質の移動速度が異なるため，小さいタンパク質が下に，大きいタンパク質が上になるように展開される．3) タンパク質を可視化するために，たとえば銀染色とよばれる方法でタンパク質を染め，このゲルをイメージスキャナで読み取る．4) 分離されたタンパク質の位置を認識ソフトウェアで解析する．この詳細は，参考文献<sup>[10][11][12]</sup>に紹介されている．この結果の例を図3に示す．横軸に電荷 (Isoelectric Point) ごと，縦軸に分子

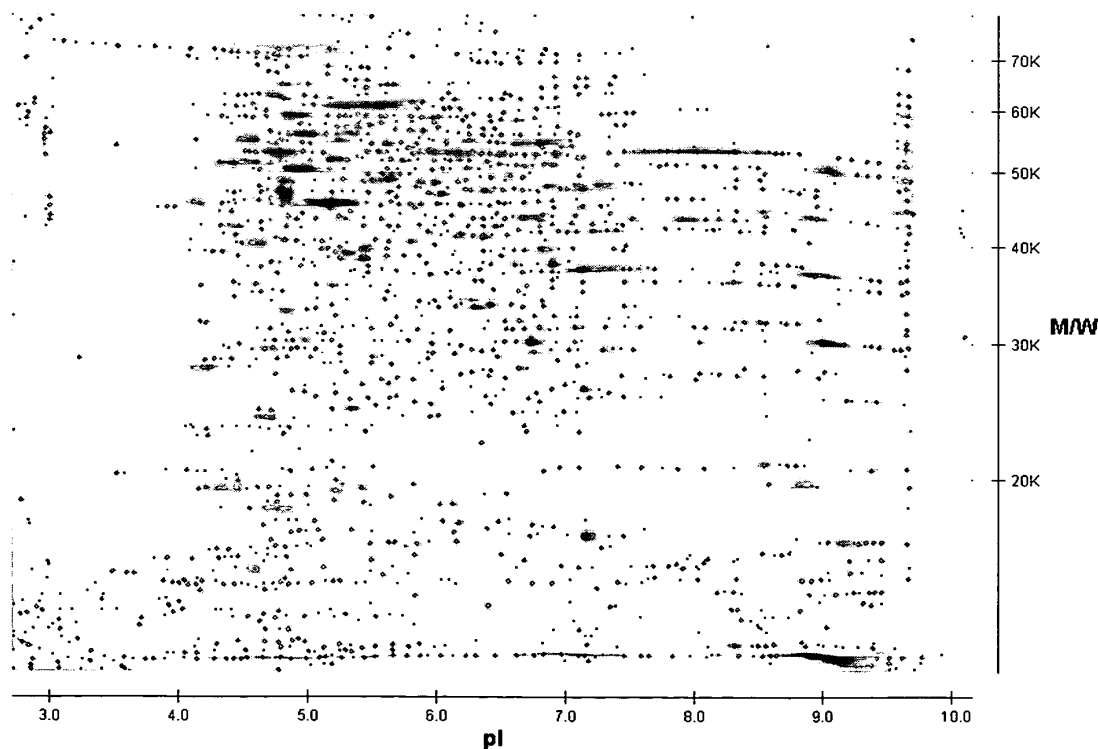


図3: C.elegans (線虫) の二次元電気泳動結果

量 (Molecular Weight) ごとのタンパク質が分離されている。黒茶色に染まった部分が分離されたタンパク質の集まりであり、これをスポットと呼ぶ。この画像を手順4)の認識ソフトウェアで処理することで、各スポットの位置が自動または手動で認識される。

解析対象となるスポットを決めると、次は、質量分析実験によりそのスポットのタンパク質の同定を行う。二次元電気泳動法により平面状ゲルに分離したタンパク質のスポットに対し、ペプチド・マス・フィンガープリント法と呼ばれる方法を使って詳細な質量を分析する。まず、各スポットをゲルから一つずつ切り出す。次に、消化酵素（トリプシンなど）を用いてタンパク質をアミノ酸配列の特定位置で切り離す。分割されたタンパク質断片を質量分析器で計測して、分割されたそれぞれの質量位置にピークがある質量スペクトルを得る。

タンパク質の同定は、既存のタンパク質データベースを用いたデータマイニングにより行う。データベース中のタンパク質に消化酵素を適用した場合の質量スペクトルを、コンピュータで計算する。この計算結果と観測された質量スペクトルが同じパターンであれば、そのタンパク質であると決定する。同じ性質を持ったタンパク質は複数存在する場合があります。そのため、複数の候補の中から研究者が似た生物種や部位などを判断し、タンパク質を同定する。

これら実験を繰り返し、研究ターゲットとなる状態から、タンパク質の機能を推定し論文などで発表する。

## 5. プロテオーム XML

プロテオミクス研究者の情報共有を容易に実現し、研究を促進させるために、XMLを利用したデータ共有フォーマットを提案する。このXMLフォーマットの特徴は、Simpson<sup>[13]</sup>ら、および、次田<sup>[14]</sup>の提案した、データベースに必要と考えられる解析結果情報に加えて、実験過程も記述することが可能である。

### 5.1. XML 記述方式

提案するXML記述はプロテオーム実験プロトコルに対応した構造を持っている。おおまかな構造を図4に示す。それぞれの項目は、4節で示した実験プロトコルに対応している。

一連の実験をまとめるルート項目“Proteome”の直下に項目“Gel”を記述する。項目“Gel”は一回の実験過程や実験結果を示す項目である。研究者は同じ試料に対して複数回の実験を行ったり、正常や疾患の違いを比較したりすることに対応するため、項目“Gel”は複数記述することができる。

項目“Gel”の下には、項目“Submitter”、“Source info.”、“Sample Preparation”、“Gel Conditions”、“Gel Image”、“Spot List”がある。

項目“Submitter”は実験者や記述者の名前などである。項目“Source Info.”は、研究ターゲットとなる試料情報であり生物種や部位の名前などである。項目“Sample Preparation”は、サンプル調整方法であり、細胞の破壊方法やタンパク質抽出に使用した試薬などの情報を記述する。項目“Gel Conditions”は、二

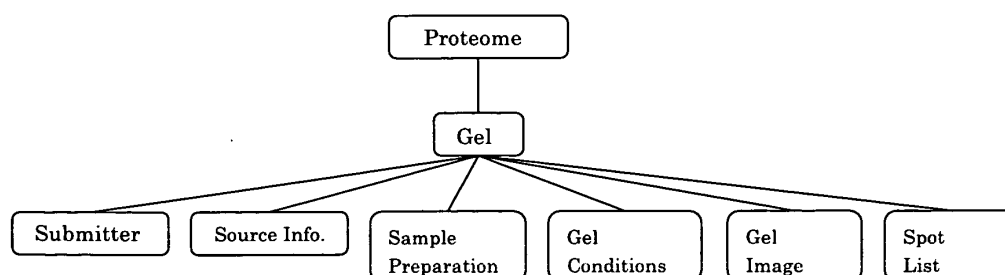


図4: プロテオーム XML の構造

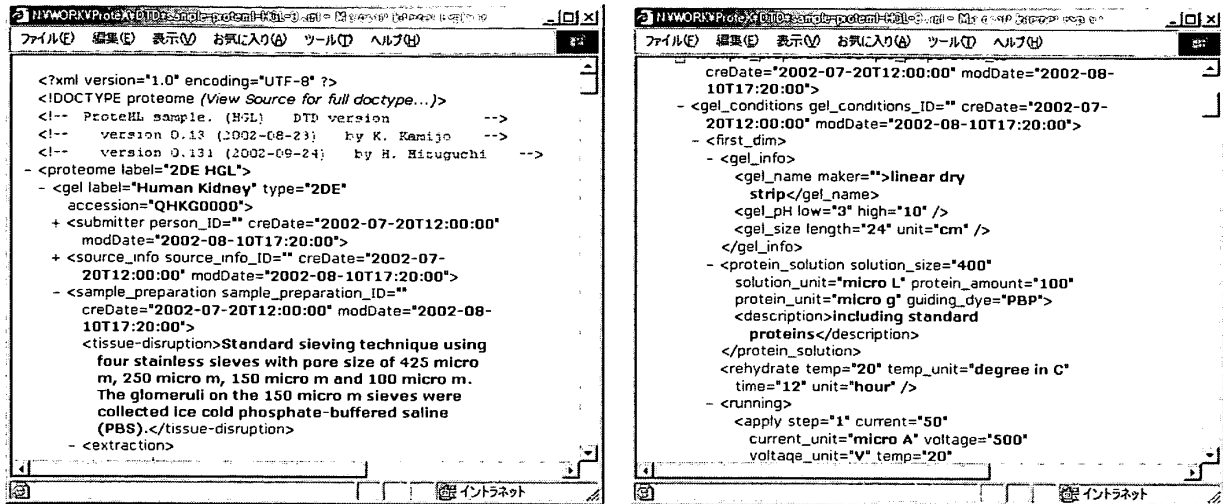


図5: “Sample Preparation” (左) と “Gel Conditions” (右) の記述例

次元電気泳動時の電圧などの条件やゲルを調整するために利用した試薬などである。項目 “Gel Image” は、画像情報であり画像サイズ、利用したイメージスキャナ名称などを記述する。項目 “Spot List” は、各スポットの情報である。この項目の下には、二次元電気泳動結果のスポット位置や、質量分析器の解析結果、タンパク質名称などを記述できる。

項目 “Sample Preparation” と項目 “Gel Conditions” の例を図5に示す。このように、我々の XML 記述方式は、他のプロテオーム記述方法と異なり、実験過程を細かく記述できる。

これらの構造は XML で記述することにより、データ交換に適した汎用性の高いフォーマットとなっている。XML はテキストファイルとしてデータを保存でき、しかも、ソフトウェアでの処理方式も決まっているため、人が見て理解でき、コンピュータ同士のデータ交換にも適している。また、データ共有に必要なデータ交換の容易さや、XSLT を利用したデータ構造の変換といった点も活用できる。

## 5.2. プロトタイプシステム

このような記述を有効に利用するためのプロトタイプを開発した。このプロトタイプは、

我々のプロテオーム XML 記述にしたがってデータを記述できるエディタと、XML データを登録検索できるデータベースサーバになる。

### プロテオーム XML エディタ

プロテオーム XML エディタは、プロテオーム実験プロトコルに対応した入力インタフェースを持ち、各種データインポート機能を持つことにより、プロテオーム XML 記述に対応したデータを容易に作成することができる。XML エディタの適用例を図6に示す。提案した XML 記述方式は、実験過程を明確に記述できるが、これにしたがって記述するためには、試薬名称、適用順序など多くの入力が必要とする。本 XML エディタは、このような入力作業を軽減することを目的としている。

本 XML エディタは実験プロトコルに従って記述することで、容易にデータを作成することができる。左側のツリービューの各項目は上から順番に実験プロトコルの各ステップを示している。ツリービューのステップを選択することで、そのステップに必要な入力項目を表示する。入力項目はタブで示され、各タブはより詳細な実験手順を示している。さらに、再利用可能なデータをエクスポートしたり、インポートしたりすることができる。たとえば、同じ生物種の同じ部位の細胞であ

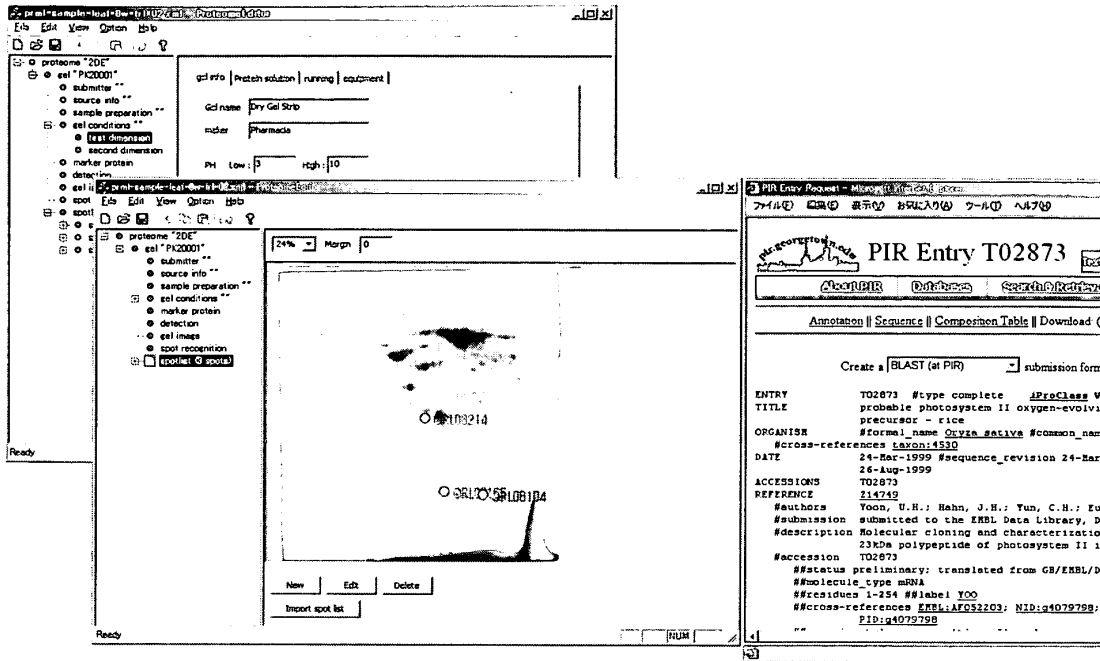


図 6: エディタ画面例

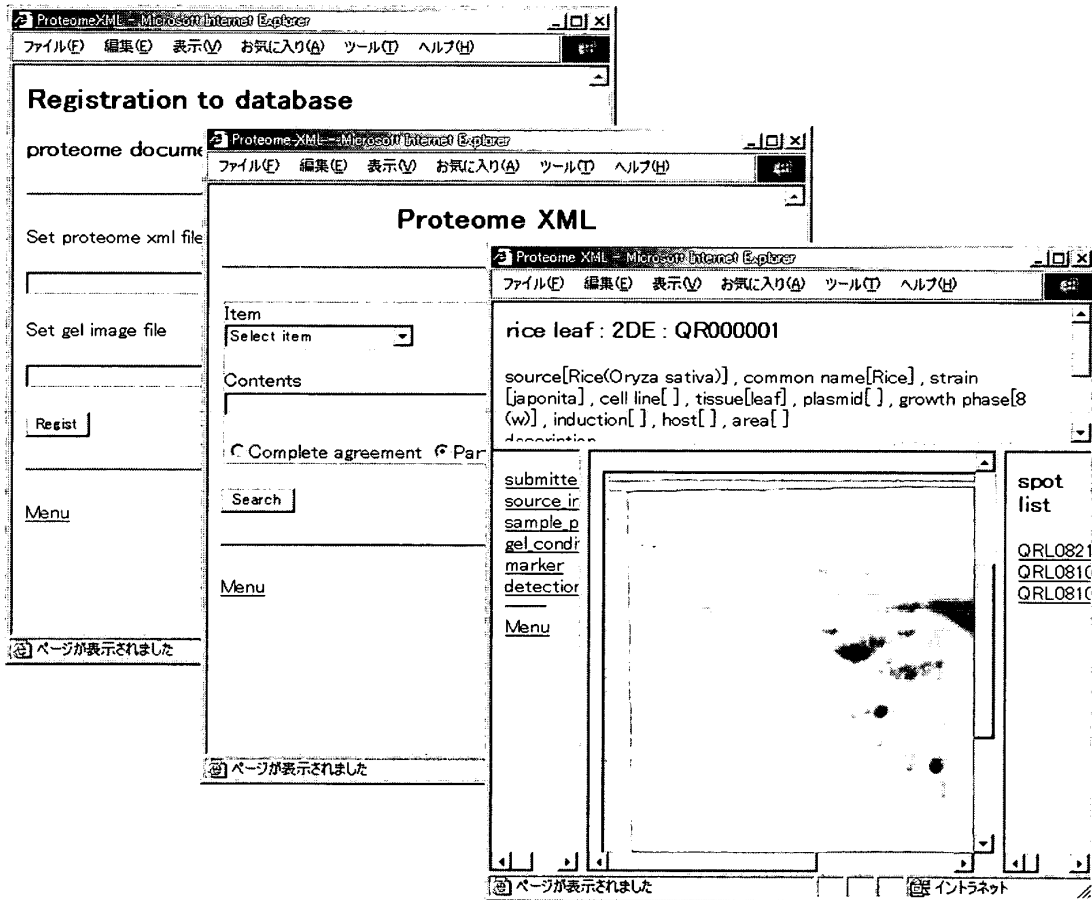


図 7: データベースサーバ例

れば、そのサンプル調整方法は同じ場合が多い。このエディタでは、一度記述したサンプル調整方法を XML 形式で保存し、再度読み出すことができる。また、質量分析器などの解析機器のデータを読み込むことで、入力の手間を軽減することができる。

さらに、二次元電気泳動結果の画像上にスポットを表示でき、各スポットのデータはデータベースへのリンクとなっている。研究者が同定したタンパク質の公開データベース名とエントリー識別子を入力し、ボタンをクリックすることでデータベースエントリーをすぐに閲覧することができる。

### XML データベースサーバ

データベースサーバは、複数のプロテオーム XML データからタンパク質の名称や、生物の名称だけでなく、利用した試薬の名前なども利用して検索することができ、必要であれば対象となる XML データを表示したり、ダウンロードしたりすることも可能である。データベースサーバのアクセス例を図 7 に示す。

試薬まで検索できることにより、同じ生物や同じ部位について研究している研究者がどのような試薬を用いてプロテオーム解析を行っているかなどを検索することもできる。これにより、実験結果の比較が可能であるかを判断することや、実験の妥当性を検証するために追試を行うといったことが、別な研究者でも可能となる。ダウンロードできることで、前述のエディタのインポート機能と共に利用すれば、別の研究者の試薬データを利用して自分の実験を行うこともできる。

## 6. まとめ

プロテオーム解析の結果だけでなく実験過程も記述できる XML 記述方式を提案し、プロトタイプシステムを作成した。XML 記述方式はサンプル調整などプロテオーム実験プロトコルを反映したものである。XML を利用することでデータ共有を容易に行うことが

できる。プロトタイプシステムは、XML 記述データを簡単に作成できるエディタ、XML 記述データを登録検索するデータベースサーバを作成した。

今後この XML 記述方式における DTD<sup>[15]</sup> や XML Schema を、プロテオーム研究者の国際組織である AOHUPO (Asia Oceania Human Proteome Organization) と共に決定し、XML エディタを配布する予定である。

## 謝辞

論文執筆にあたり、ご助言をいただいた NEC インターネットシステム研究所の久寿居大氏に感謝いたします。

## 参考文献

- [1] 日本 DNA データバンク: DDBJ の統計「DDBJ/EMBL/GenBank database 塩基配列の推移」, <http://www.ddbj.nig.ac.jp/ddbjnew/statistics-j.html> (2002 年 10 月 25 日参照).
- [2] National Biomedical Research Foundation, Georgetown University Medical Center: "PIR(Protein Information Resource)", <http://pir.georgetown.edu/> (2002 年 10 月 25 日参照).
- [3] Swiss Institute of Bioinformatics: "SWISS-PROT Protein knowledge base", <http://tw.expasy.org/sprot/> (2002 年 10 月 25 日参照).
- [4] National Center for Biotechnology Information: "XML at NCBI", <http://www.ncbi.nlm.nih.gov/IEB/ToolBox/XML/> (2002 年 10 月 25 日参照).
- [5] European Bioinformatics Institute: "The XEMBL Project", <http://www.ebi.ac.uk/xembl/> (2002 年 10 月 25 日参照).

- [6] 日本 DNA データバンク: “XML Central of DDBJ”, <http://xml.nig.ac.jp/index.html> (2002年10月25日参照).
- [7] LabBook inc.: “BSML (Bioinformatic Sequence Markup Language)”, <http://www.bsml.org/> (2002年10月25日参照).
- [8] Proteometrics, LLC and Proteometrics Canada Ltd: “BIOML (BIOpolymer Markup Language)”, <http://www.bioml.com/BIOML/index.html> (2002年10月25日参照).
- [9] Microarray Gene Expression Data Society: “MAGE-ML (Microarray Gene Expression - Markup Language)”, <http://www.mged.org/Workgroups/MAGE/mage.html> (2002年10月25日参照).
- [10] 磯辺敏明, 高橋信明 編: 「プロテオーム解析法 タンパク質発現・機能解析の先端技術とゲノム医薬・創薬研究」, 羊土社, 282p, 2000.
- [11] 岡田雅人, 宮崎香 編: 「改訂 タンパク質実験ノート 上 抽出と分離精製」, 羊土社, 191p, 2002.
- [12] 岡田雅人, 宮崎香 編: 「改訂 タンパク質実験ノート 下 分離同定から一次構造の決定まで」, 羊土社, 163p, 2002.
- [13] Simpson, R.J.; Tsugita, A.; Celis, J.E.; Garrels, J.I.; Mews, H.W. : “Workshop on two-dimensional gel protein databases”, *Electrophoresis* 13, pp.1055-1061, 1992.
- [14] 次田皓: 「ポストゲノムにおけるプロテオーム研究」, 情報知識学会誌, Vol.10, No.4, pp.32-42, 2001.
- [15] Kamiyo, K.; Yamazaki, T.; Tsugita, A: “A Proposition of XML Format for Proteomics Database”, *Proc. Of 18th International CODATA Conference*, p.50, 2002.

(2002年11月14日受付)  
(2002年11月30日採録)



健診情報のための電子的交換規約  
— Health Data Markup Language (HDML) —  
Electronic Data Interchange Protocol for Healthcare  
— Health Data Markup Language (HDML) —

原 正一郎<sup>(1)</sup>, 杉森 裕樹<sup>(2)</sup>, 古海 勝彦<sup>(3)</sup>,  
東福寺 幾夫<sup>(4)</sup>, 窪寺 健<sup>(4)</sup>, 河合 正樹<sup>(4)</sup>, 吉田 勝美<sup>(2)</sup>  
Shoichiro HARA<sup>(1)</sup>, Hiroki SUGIMORI<sup>(2)</sup>, Katsuhiko FURUMI<sup>(3)</sup>,  
Ikuo TOFUKUJI<sup>(4)</sup>, Takeshi KUBODERA<sup>(4)</sup>,  
Masaki KAWAI<sup>(4)</sup> and Katsumi YOSHIDA<sup>(2)</sup>

健康審査（健診）には多数の施設が関わっており、これらの施設の健診情報システム間でデータを交換する必要がある。ところが健診に関わるデータ構造は多様であり、健診依頼元へ検査結果を電子的に還元することすら容易ではない。そこで平成8年より、日本総合健診医学会情報委員会は、健診データの有効利用を図るために日本保健福祉医療情報システム工業会と合同委員会を組織し、標準健診データ交換規約 (HDML: Health Data Markup Language) とツールの開発に着手した。HDMLはSGML/XMLを基礎とし、既存の医療情報交換規約との互換性を考慮しつつ、健診に特化したデータ交換規約である。本稿ではHDMLの構造と、HDMLツールを用いた評価試験の結果について述べる。HDMLの導入によりデータ交換にかかわるコストの削減が可能となる。さらにデータ交換の段階でデータ構造が標準化されるため、複数の医療施設で発生した健診データを生涯健康管理データベースとして集積・管理できるようになり、生活習慣病などの予防医学にも大きく貢献することが期待される。

A lot of healthcare facilities should be concerned with health-checkups, and healthcare data must be exchanged among information systems of these facilities. However, as healthcare data structure is heterogeneous among facilities, even the electronic data transfer of test results to the requester is not easy. The Joint Working Group of JMHTS (Japan Society of Multiphasic Health Testing and Service) and JAHIS (Japanese Association of Healthcare Information Systems Industry) commenced developing standard healthcare data exchange protocol (HDML: Health Data Markup Language) and tools for effective data utilization among healthcare facilities. HDML is SGML/XML-based data description protocol, with considering compatibility with existing standard medical/health information exchange protocols, that specializes in healthcare data exchange. The structure of HDML and the results of the feasibility study with HDML tools are described. HDML can reduce the cost related to data exchange. In addition, HDML is expected to contribute to a preventive medicine such as lifestyle diseases greatly because all data structure is standardized during data exchanging and all healthcare data generated by different medical/health facilities will be easily accumulated and managed as "lifetime healthcare data management system."

キーワード：XML, SGML, HDML, Standardization, 標準化, Health Checkups, 健康診断

(1) 国文学研究資料館研究情報部, Research Information Department, National Institute of Japanese Literature

(2) 聖マリアンナ医科大学予防医学教室, Department of Preventive Medicine, St. Marianna University School of Medicine

(3) 産業医科大学医学部医療科学講座, Health Care System Administration, School of Medicine, University of Occupational and Environmental Health

(4) JAHIS 健診情報委員会, Committee of JAHIS, Japan

## 1. 概要

日本の労働安全衛生法では健康審査（以下では健診）の受診を国民に義務づけている。健診データは定期的かつ全国規模で収集されるため、疾病の早期発見と早期治療（二次予防）のための重要なツールとなっている。また近年の急速な高齢者人口の増加に伴い、健診の役割は二次予防から、地域における疾病予防と健康増進（一次予防）に移りつつある。さらに、日本の疾病構造が伝染病から生活習慣病を主体としたものに変化したため、医療費が増大し、国の医療福祉予算を圧迫しつつある。このような中で、多くの保健医療従事者は、個人の生活環境に適したライフスタイルの改善が生活習慣病を効率的に抑制し、ひいては医療費の減少を実現する重要な方法であると考えられるようになった。そのためには継続的な健康状態のモニタリングが必要であり、健診は実質的に唯一の手段である。したがって健診データを基盤とした生涯健康情報管理システムのような仕組みを構築することは必要不可欠であると考えられる。

健診情報の管理は、閉鎖的なシステム環境下、例えばデータが特定の健診施設内や企業内において収集・蓄積・閲覧される場合は容易である。しかし健診には多数の健診依頼元と健診施設が関わっており、複数の健診情報システム間で健診データを交換・共有する必要がある。ところが健診に関わるデータの構造は多様であり、健診依頼元へ検査結果を電子的に還元することすら容易ではない。そこで平成8年より、日本総合健診医学会情報委員会 (JMHTS : Japan Society of Multiphasic Health Testing and Service) は、標準化による健診データの有効利用を図るため、日本保健福祉医療情報システム工業会 (JAHIS: Japanese Association of Healthcare Information Systems Industry) と合同委員会（以下では合同委員会）を組織した<sup>[1]</sup>。標準化の目的は、各健診情報システムに分散している健診データの共有を実現し、これらのデータに基づいて個人の健康状態を正確に評価・予測

することにある。

合同委員会では健診データの標準化に対するニーズを探るため、平成8年度にアンケートによる実態調査を行った<sup>[2]</sup>。調査には健診施設60施設、健診依頼元13施設が参加した。調査の結果、健診施設の73%、健診依頼元の77%が、健診データの標準化について必要性を訴え、一刻も早い方策を希求していた。

そこで合同委員会では、健診データの標準化に向けて、(1) データモデルの検討、(2) 標準用語集の作成、(3) 符号化法の3つの検討を行った。次に検討結果に基づいて、健診情報システム間で健診データを交換するための標準健診データ記述規約 (HDML: Health Data Markup Language) を定義した<sup>[2][3][4]</sup> (なお、本稿のHDMLはThe World Wide Web Consortium (W3C) から提案されている Handheld Device Markup Language (URL: <http://www.w3.org/TR/NOTE-Submission-HDML.html> など) とは異なる規約である)。HDMLはISO標準の文書記述言語であるSGML (Standard Generalized Markup Language: ISO 8879:1986 および JIS X 4151-1992)<sup>[5][6]</sup> およびXML (eXtensible Markup Language)<sup>[7]</sup> ベースの符号化言語である。HDMLはHL7 (Health Level 7)<sup>[8]</sup> のOBXセグメントを参考に合同委員会独自の拡張を施したものであり、HL7のサブセットと位置づけることもできる。HDMLの特色は、

- ① 健診に特化した簡易な規約であり、実装や修正が容易であること
- ② 他の規約との互換性を考慮していること
- ③ SGML/XML ベースであるためデータの記述能力が高いこと

にある。HDMLでは、病歴・所見・臨床検査・診断結果など一般的な健診データ項目を組織化するためのDTD (Document/Data Type Definition) を定義している。さらにコンピュータによるデータ解釈を支援するために、項目名・検査方法・単位・検査装置・標準値などの属性情報の種類と記述法も定義している。

HDMLの有効性を検証するために、合同委

員会では個別の健診情報システムに蓄積されている健診データと HDML データを相互に変換する HDML 変換ツールを試作して評価試験を行った。試験の結果、HDML 変換ツールは送信側と受信側の情報システムの間で殆どの検査値を正確に交換することができた。

本稿では HDML の開発について、まず第 2 章で健診データの標準化を進めた背景と標準化の要件について考察する。第 3 章では標準化作業における検討事項について述べ、第 4 章では HDML の詳細について説明する。第 5 章では HDML 変換ツールの概要、第 6 章では評価試験について述べ、最後に HDML の課題について考察する。

## 2. 標準健診データ交換規約の要請

本章では、健診データの標準が必要な背景と、標準化されたデータ交換規約が持つべき条件について述べる。

### 2.1. 健診データ交換規約の要件

保健医療データの標準化は欧米で特に盛んであり、臨床データ伝送規格 (ASTM E1238 : 以下 ASTM と略す)<sup>[9]</sup>、会計処理・入院管理・臨床検査などのための伝送規格 (HL7)、あるいはヨーロッパ連合の CEN/TC251 (The European Committee for Standardization (CEN): Technical Committee 251) による規格化の推進などが挙げられる。国内でも ASTM の適用例、医療情報システム開発センターによる臨床検査データ交換規約 (暫定版)<sup>[10]</sup>の作成などがあるが、全体として動きが遅かった。近年、JAHIS を中心として保健医療情報システムに関連する様々な規約の調査あるいは策定作業が行われている。

標準健診データ交換規約の目的は、健診依頼元や健診施設に眠っている健診データを効率的に交換する手段を提供し、健診データの高度利用による健康増進の実現に寄与することである。このような状況における合同委員

会の目的は、健診情報システム間における健診データの標準化された交換法を実現するための基礎的な作業として、データ項目および用語の選定と情報モデルの構築であった。

この保健医療データの基本は紙というメディア上に記録されている文字データ (いわゆるカルテ) であり、電子カルテの主要な目的は既存のカルテのデータ構造を電子メディア上に再現することである。そこで ASTM や HL7、あるいは他分野の規約であるが EDIFACT など、現在公開されている幾つかのデータ交換規約を調べてみると、これらのデータ構造は樹形図とほぼ等価であることが分かる。これは文脈自由文法のクラスに相当する。合同委員会を対象とする健診データの構造も、調査票や個人票が基礎であるため、やはり樹形図で表現できる。したがって健診データ交換規約のデータ記述能力も文脈自由文法と同等以上のクラスでなければならない。

健診データを効率的に収集できれば、これらのデータを適切に処理して個人の健康状態を評価・予測し、適切な健康管理・指導を行うことが期待できる。さらに疫学的調査に適用すれば、集団の特性を地域的・時系列的に明確化し、職域・地域における健康活動の指針や手段を決定することなども可能となろう。

健診で生成されたデータをこのような目的で利用するためには、データの値に加えてデータの解釈に必要な情報が必要である。例えば赤血球数の単位は、検査会社より「万/mm<sup>3</sup>」あるいは「10<sup>6</sup>/μL」などと異なっている。また血清コレステロール値が「220 mg/dL」であった場合、これは多くの健診において正常域と高コレステロール血症の判定境界値である。ところが血清コレステロールの正常域は検査法により微妙に異なるので、正確な比較や判断を行うためには検査法の情報が不可欠である。この他にも、「値」の解釈に必要な属性情報として、計測装置・単位・標準範囲・データ型などが挙げられる。これらの属性は、受信側の健診情報システムがデータを解釈して、受信側のデータ構造に自動的に変換する

ために必要な情報である。健診データには所見や病歴などの定性的データも含まれる。これらの多くはコード化されているため、定性データの属性情報としてはコード体系名なども必要である。また、医師・保健婦あるいは栄養士らによるコメントやメモなどの文書も属性情報に含まれる。

ASTM や HL7 などの標準規約においても、計測値に加えて属性情報が記述できるようになっている。しかし HDML が対象とする健診に必要なデータ項目や属性情報の範囲と、HL7 などの臨床用規約の対象範囲は必ずしも一致していない。また ASTM や HL7 は国際的かつ大規模な規約であるため、その改訂や修正には時間と手間がかかる。そこで、既存の標準規約との互換性を考慮しつつ、健診データ交換に特化し、かつ改訂や修正が容易で軽装な標準規約が必要となった。

## 2.2. 標準化の要件

ここで、健診データの標準化の考え方について整理する。まず健診データの現状は以下のように要約される。

- ① 健診データの多くは、健診施設などのコンピュータに蓄積されており、個別には統計処理などが可能である
- ② 個人の健康管理を適切に行うには、全てのデータソースからデータを収集できなければならないが、実現していない
- ③ 集団の健康状態を正確に把握して適切な政策決定を行うには、分散しているデータを集めて情報に変換しなければならない。健診結果を市町村単位で集計し報告する制度もあるが、全データが電子化されているわけではなく、集計法も標準化（検査法、正常範囲、単位、あるいは記述法）されているとは言い難く、疫学調査に耐えうるデータの質とは言い難い

一方、技術的な面からみると、以下のような問題点が指摘される。

- ① データの大部分は各コンピュータに蓄積され、データを交換・処理する技術は既に確立している
- ② 各健診情報システム用に大規模なプログラム開発を必要とせず、データの交換が可能なデータ構成および記述法についての標準規格は存在してない。したがって健診情報システム間における効率的なデータ交換は実現していない。一方、特定の健康情報システム間におけるデータ交換を実現しようとする、高額の投資が必要となる

標準健診データ交換規約は上記②の解決を目指したデータ記述法を規定し、健診データの高度利用を実現することにある。しかし標準化には以下の問題点も指摘されている<sup>[11]</sup>。

- ① 情報システムは本質的に heterogeneous である。例えば、薬局のコンピュータには薬歴情報、検査会社のコンピュータには臨床検査情報、病院のレセプトコンピュータには処方と診断情報が蓄積されているが、これらのデータ構造は各々異なっている。これらのデータ構造を統合することは非現実的である。仮に統合できたとしても、時間がかかる上に、データ項目の変更などに伴う全システムの修正コストなど問題が多い
- ② 過度の標準化は、発展の原動力となりうる多様性を阻害する
- ③ 標準化には時間がかかる。システム開発企業は標準化が終了するまで開発を止めておけない。その結果、可般性のある製品を開発できず、特定利用者向けの製品しか供給できない。このような製品は高価でライフサイクルが短く互換性がないので、一般利用者にとっては使いにくいシステムとなる。一方、標準化が完了したときには、それ自体が時代遅れになっている可能性がある

以上の点を考慮し、標準健診データ交換規約の定義作業においては以下の点に留意した。

- ① 特定の蓄積メディア、物理的・論理的データフォーマット、ネットワークなどの転送方式、データ項目、データ配列、データ属性、情報システムのハードウェアおよびアプリケーションなどには依存しない規約とする
- ② 対象を健診に特化して、電子カルテのように過度に複雑なデータは扱わない
- ③ 健診情報システム間でデータ交換を行う際のデータ記述に関する軽装な規約とする。これにより、規約の策定ならびに改訂・修正を短期間で行えるようにする
- ④ 交換されたデータを解釈するうえで必要なデータ構造と属性情報の記述法のみを規定する。これにより送信側は、受信側のデータ構造を考慮する必要がなくなる。受信側は受け取ったデータを属性情報に基づいて解釈し、受信側システムのデータ構造に変換して蓄積・処理する。不要なデータは廃棄しても構わない。これにより、交換データを受信側システムの目的に最適化した形で処理することが可能となる
- ⑤ アプリケーションなどのデータ処理法は対象外とする
- ⑥ データは可読テキストとする。これは、データ処理と可搬性の実現を容易にするための措置である
- ⑦ 国際的な保健医療データ交換規約との互換性を十分に考慮する
- ⑧ 紙に書かれている文書が基本であるため、

データ記述レベルを文脈自由文法のレベルとする

- ⑨ 標準健診データ交換規約は OSI 参照モデルの第7層以上に相当する。したがって、データの転送媒体・経路制御・セキュリティなどは考慮の対象外とする

### 2.3. 標準健診データ交換規約導入のシナリオ

標準健診データ交換規約を導入するシナリオとしては以下の2つを想定している [12]。

#### (1) 施設の健診データを他施設に伝送する

ある企業（健診依頼元）が職場の定期健診を外部の健診施設（総合健診施設など）に委託している場合、健診依頼元と健診施設の健診情報システムが異なるため、依頼元に健診データを電子的に返却するには手間と時間がかかる。また健診施設が複数の依頼元と契約している場合、依頼元のシステムごとに健診データを加工する必要がある。もし健診依頼元が  $N$  件あった場合、健診依頼元のデータを健診施設用に変換するために  $N$  本、健診施設のデータを健診依頼元用に変換するために  $N$  本、つまり健診施設は  $2N$  本の変換プログラムを用意し、維持・管理しなければならない。もし標準健診データ交換規約によるデータ交換ができれば、健診施設は自己のデータ構造と標準規約の間の相互変換用に2本のプログラムを用意するだけですむ（図1）。

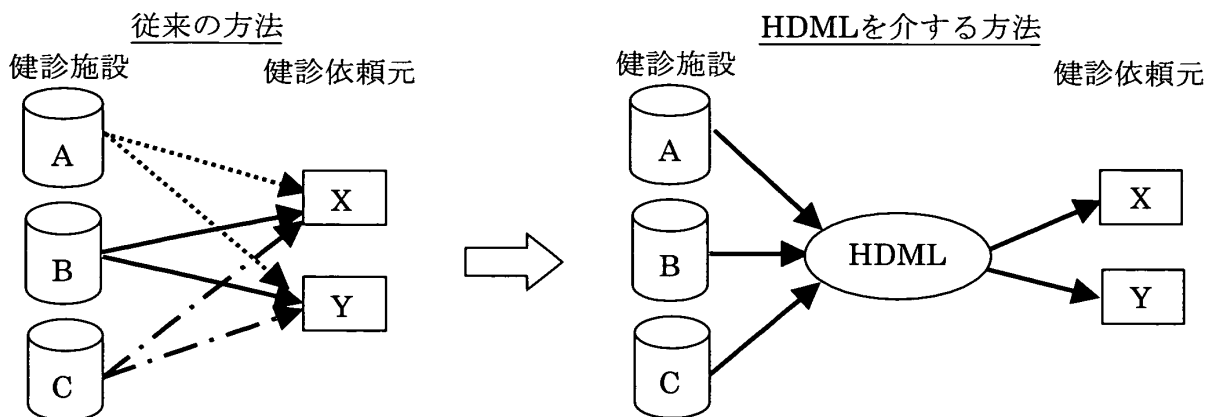


図1: 健診データを他施設に伝送する方法

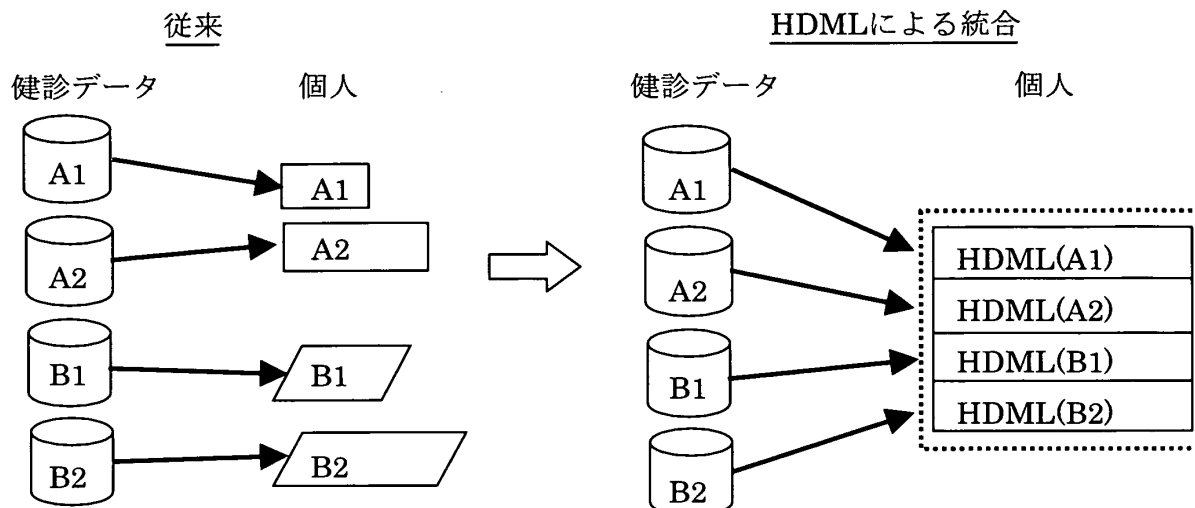


図 2: 健診データを統合する方法

## (2) 個人の健診データを統合する

健診受診者が転勤などの理由により複数の健診施設で受診した場合、健診依頼元では健診施設ごとに異なったデータ構造を依頼元のデータ構造に変換しなければならない。この場合でも標準健診データ交換規約を介することで、異なる健診情報を一括管理することが可能である(図2)。また同じ方法により、ライフステージが移行(学校から職域、職域から地域など)しても個人を軸とした生涯健康情報管理データベースの構築も可能である。

## 2.4. SGML/XML の導入

標準健診データ交換規約が独自のシンタクスを採用した場合、パーサ(parser)などのツールは全て自前で開発しなければならない。一方、標準シンタクスを利用すれば、システム開発やデータの維持・管理は比較的容易となる。そこで標準健診データ交換規約ではSGMLをシンタクスとして採用した(現在、XMLへの移行を進めている)。SGML/XMLには、パーサやエディタなどのツールが市販されており、これらを組み合わせることにより開発効率を上げることができる。以下に標準健診データ交換規約の策定においてSGML/XMLを採用した理由をまとめる。

- ① SGML/XMLは文脈自由文法のクラスに属するデータ記述言語であり、樹形図として表現できるデータ構造を記述できる。保健医療データの多くも樹系図として表現でき、データ記述上の親和性が高い
- ② SGML/XMLは可読データであり、ハードウェアやソフトウェアから独立してデータの交換や保存を行うことができる
- ③ SGML/XMLパーサはデータ要素の出現順序・出現回数などを検証する。これによりデータの構造的誤りを事前に防ぐことができる。SGML/XMLパーサはDTDからコンパイラを用いて生成できる
- ④ SGML/XMLでは属性を利用してデータに多様な情動的修飾(例えばデータ間の関連付けを示すリンク情報など)を施すことができる。またHTMLはSGML/XMLのシンタクスを利用しているので、SGMLデータからHTMLへの変換は容易である

## 3. 標準健診データ交換規約の実現

本章では、標準健診データ交換規約の構築に際して、合同委員会が実施した作業や検討事項について述べる。

| 分類 | JAHIS標準用語 | JAHISコード | 臨床検査項目分類 | JLCA10コード | 類義語                |
|----|-----------|----------|----------|-----------|--------------------|
| 血液 | 白血球数      | 300      | 白血球数     | 2A010     | WBC                |
|    | 赤血球数      | 301      | 赤血球数     | 2A020     | RBC                |
|    | 血色素量      | 302      | 赤血球数     | 2A030     | ヘモグロビン,<br>Hb, HGB |
|    | ヘマトクリット   | 303      | 赤血球数     | 2A040     | Ht, HCT            |
|    | 血小板数      | 304      | 血小板      | 2A050     | PLT                |

図 3: JAHIS 類義語表

### 3.1. 語彙

各健診情報システムで使用されているデータ項目名や値の中に現れる語彙はシステムごとに異なっている。そのため、データ交換を実現するためには語彙の統一が不可欠であった。合同委員会の議論において、語彙は少なくとも4つのタイプに分類された [2][13]。

第一のタイプは名前、アドレス、IDなどの「命名的な語彙」である。命名的な語彙の場合、語彙間の対照表(mapping tableあるいはthesaurus)を利用することにより、各健診情報システムで利用している語彙をデータ交換で使用する語彙と相互に変換することは容易である。合同委員会では健診情報システムで利用されている語彙を収集するため、JAHISの会員企業を中心にアンケート調査を行った。収集された語彙は同義語ごとにグループ分けされ、各同義語グループには見出し語とコードを付け、これを標準健診データ交換規約で利用する語彙とした(以下では標準語彙)。標準語彙は医療情報システム開発センターから発行された「健康医療カードシステム標準化マニュアル 第一版」[14]に従い、そこに記載されていない語彙については、該当する医学会等から発行された用語集によった(図3)。

第二のタイプはICD(International Classification of Diseases)[15]やSNOMED(Systemized Nomenclature of Medicine)[16]などで利用されている「標準コード」である。この場合、データの送信側と受信側の健診情報

システムが同じ標準コードを利用していれば問題はない。また異なった標準コードを利用している場合でも、標準コード間の変換テーブルあるいは変換規則を持っていれば、このタイプの語彙の変換は比較的容易である[17][27]。

第三のタイプは「ローカルコード」である。ある質問に対する肯定の記述法が、健診情報システムごとに「y」、「1(いち)」あるいは「はい」などと異なっている場合である。また「性別」を「1」と「2」あるいは「m」と「f」で表すような場合もここに含まれる。この場合もローカルコード間の変換テーブルあるいは変換規則があれば、相互変換することは可能である。

第四のタイプは「ローカルな意味コード」である。あるシステムでは「痛み」を5段階に分類しているが、別のシステムでは2段階のみという場合である。このようなデータを相互に変換する方法は未だ実現されていない。

合同委員会では、標準健診データ交換規約が対象とする語彙体系を1, 2, および3に限定している。

### 3.2. データモデル

データモデルは語彙の意味を分類するために導入された[2]。データモデルは「一般機能」と「サブシステム」から構成されている。一般機能は臨床的機能、例えば神経機能、筋肉骨格機能、全身機能、炎症などからなる。サ

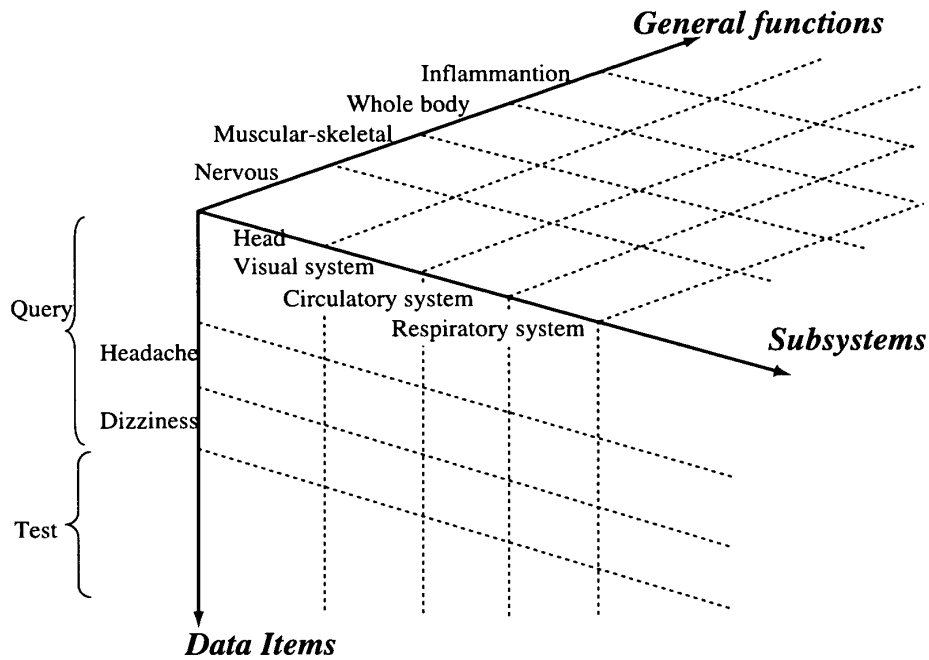


図 4: データモデル

ブシステムは解剖学的サブシステム，例えば循環器系サブシステム，呼吸器系サブシステムなどからなる．一般機能とサブシステムは2次元のマトリクスとして表現され，各データ項目はマトリクスの上の適切な位置に割り当てられる（図4）．このモデルにおいてデータ値は

値 = (一般機能，サブシステム，データ項目名)

で表される．例えば、「頭痛」は様々な場面で表れるデータ項目である．頭痛というデータ項目は，ある場面では脳血管系の診断で利用され，別の場面では視覚系の診断に利用されることもある．したがって，同じデータ項目であっても，その目的や意味合いをデータ項目に反映させる必要があると考えた．

ところでデータモデルは語彙を整理する段階で考え出されたものである．しかし公衆衛生あるいは疫学研究者の一部からは，データを収集した目的もデータ解析上の重要な情報であるとの意見があり，標準健診データ交換規約においても，後述するように，データモデルと同等の構造を記述できるようにした．

### 3.3. 符号化スキーム

データの値だけではなく関連する属性も交換できるようにするには，明確なデータ構造を定義する必要がある．HL7の場合，

```
OBX|100|ST|84295^Na||150|mmol/l
|136-148|H||A|F
```

のように特殊なセパレータで区切られ順序も固定された幾つかのフィールドからなるデータセグメント構造によりデータを記述する．

HL7のデータ構造は明確に定義されているので記述に曖昧さはない．その反面，必要に応じてフィールドを修正すること（追加あるいは削除）は容易ではない．さらに，データ値に関する全ての属性を予め定義することは困難であるから，必要最小限の属性定義から始めて，次第に属性情報の取捨選択を行える，柔軟な符号化スキームが必要である．合同委員会では符号化のスキームとしてSGML/XMLを採用し，これに基づいて標準健康データ交換規約HDML (Health Data Markup Language) を定義した（以下ではJAHIS健診データ交換規約 Ver.1) [2][3][4]．



## 4. HDML の構造

現時点の HDML には未定義の部分が多い。HDML の基本的構造は、メッセージ全体に関わる情報を記述したヘッダ (Header)、健診データ本体 (Body)、およびシステムによる後処理用の情報を記述したトレイラ (Trailer) から構成される (図 5)。

HDML は以下のように定義されており、

```
<!ELEMENT HDML(Header,Body,Trailer?)+>
```

“Header” と “Body” は必須であるが “Trailer” は必ずしも必要ではない。以下では HDML の特徴的な部分について説明する。

### 4.1. ヘッダ

ヘッダには、データの送信者 (Sender)、受信者 (Receiver)、健診の概要 (Project)、健診依頼者 (Client)、健診施設 (Contractor) およびデータのデフォルト情報 (Defaults) が記載される。

```
<!ELEMENT Header(Sender?,Receiver?,Project?,Client?,Contractor?,Defaults?)>
```

前述のように、検査データには正常範囲・検査法・単位などの属性が必要であり、HDML では各受診者の検査項目ごとに属性を細かく指定できる。しかし集団健診データを一括交換するような場合、データ中の各検査項目の属性は一般的に同じである。例えば、ある集団健診においてコレステロールに関する属性が受診者ごとに異なることは殆どない。もし各受診者の検査項目ごとに属性を記述すると、冗長なうえに容量も増えて、データ交換上の障害となる。そこで HDML では、検査項目ごとに共通の属性をヘッダの “Defaults” 部分に集めてデフォルト情報とした。デフォルト情報と異なる属性のみをデータ本体の “Body” 部分に記述することにより、データ交換におけるオーバーヘッドの軽減を図った。

健診データで必要な属性としては、特に HL7 の OBX (Observation/result) セグメントを参

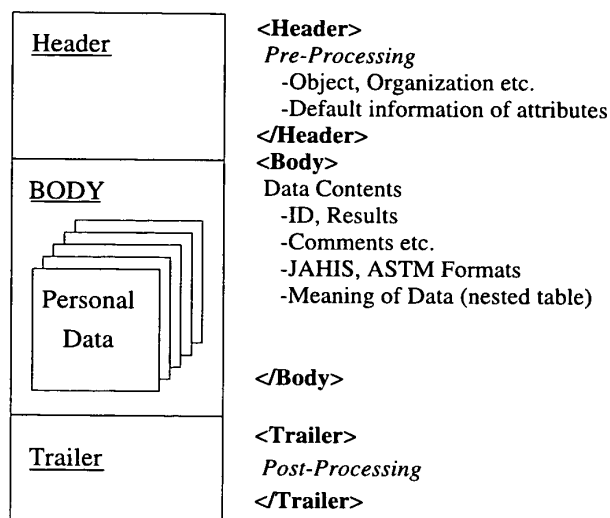


図 5: HDML の構造

考とした。さらに合同委員会では健診機関が発行する帳票類を調査した。その結果、検査データでは計測値に加えて、項目名・単位・正常範囲等が必須であることなどが明らかになった。また帳票のヘッダ部分などからは、健診機関に関する情報、検査日などを読みとることもできた。また合同委員会による検討過程で、データの疫学的利用あるいは詳細な解析を行うためには、目的・検査法・分析条件・検査装置名なども記述すべきであるとの意見が提出された。さらにレコードアクセスにも利用者に応じた制限を設定する必要があるとの指摘もあった。これらの意見を総合した結果、HDML における検査データのデフォルト情報は図 6 のようになった。

検査データを例に説明する。検査データの属性としてはデータ項目名を一意に示す項目コード “c” のみが必須である。言い換えれば、属性 “c” と検査値のみで最小限のデータを構成することができる。

属性 “CodeType” によりデータ項目名を示すコード系を指定する。合同委員会が定義した項目コードをデフォルトとしているが、他規約のコードを使用することも可能である。

```
CodeType="JAHIS"
```

属性 “DataType” は値のデータ型を表す。もしデータが定量的な検査値 (数値型) であれば、

|                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|----------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>&lt;!ATTLIST fTest</p>                    | <p>sOrder CDATA #IMPLIED ← 検査依頼番号</p> <p>c CDATA #REQUIRED ← 項目コード</p> <p>n CDATA #IMPLIED ← 項目名</p> <p>Segment CDATA #IMPLIED ← 分画番号</p> <p>Unit CDATA #IMPLIED ← 単位</p> <p>Upper CDATA #IMPLIED ← 正常範囲の上限</p> <p>Lower CDATA #IMPLIED ← 正常範囲の下限</p> <p>Decision CDATA #IMPLIED ← 検査判定</p> <p>Belief NUMBER #IMPLIED ← 確信度</p> <p>DecisionBase CDATA #IMPLIED ← 判定基準</p> <p>Status CDATA #IMPLIED ← 結果状態</p> <p>dateModified CDATA #IMPLIED ← 最終变更日期</p>                             |
| <p>ASTM<br/>Compatibility</p>                | <p>DataFormat (JAHIS  ASTM4  Local) "JAHIS" ← データフォーマット</p> <p>Method CDATA #IMPLIED ← 検査法</p> <p>Condition CDATA #IMPLIED ← 分析条件</p> <p>Equipment CDATA #IMPLIED ← 検査機器</p> <p>%atSampled; ← 検体採取についての情報</p> <p>%atOrdered; ← 検査依頼についての情報</p> <p>%atExamined; ← 検査機関についての情報</p> <p>Issued CDATA #IMPLIED ← 結果の発行日</p> <p>%atPromoter; ← 健康診断の主催者情報</p> <p>%atContractor; ← 健康診断の請負者情報</p> <p>DataType %adDataType; #IMPLIED ← データ型</p> <p>CodeType %adCodeType; #IMPLIED ← コード体系</p> |
| <p>Value<br/>Added</p>                       | <p>%atID; ← Tag ID for Link</p> <p>%atTimeStamp; ← この情報のタイムスタンプ</p> <p>%atSecurity; ← この情報の公開レベル</p> <p>%atPasswd; ← この情報のパスワード</p>                                                                                                                                                                                                                                                                                                                                                   |
| <p>Access &amp;<br/>Security<br/>Control</p> | <p>Memo CDATA #IMPLIED ← メモ</p> <p>--中略--</p>                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| <p>&gt;</p>                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

図 6: 検査データの属性情報

DataType = "MN"

と記述される。この場合、少なくとも定量値に関連した属性、つまり検査値の単位“Unit”，正常値の上限値“Upper”と下限値“Lower”は記述されなければならない。

もし“DataType”が定性型(例えばコード型)であれば、

DataType = "CE"

となる。

なおデフォルト情報を示すタグ“fTest”はEMPTYタグとして扱った。これにより、デフォルト情報は必要なデータ項目を任意の順序で記述できる。検査データのデフォルト情報の記述例を図7に示す。ここで、“CodeType”の“L”はコード体系がローカルであることを示している。

```

<Defaults>
  <fTestc="101" n="体重" Unit="Kg" DataType="NM" CodeType="L"/>
  <fTestc="100" n="身長" Unit="cm" DataType="NM" CodeType="L"/>
  <fTestc="210" n="最大血圧" Unit="mmHg" Upper="140" Lower="90" DataType="NM" CodeType="L"/>
  <fTestc="211" n="最小血圧" Unit="mmHg" Upper="90" Lower="60" DataType="NM" CodeType="L"/>
  <fTestc="400" n="血糖" Unit="mg/dl" Upper="110" Lower="70" DataType="NM" CodeType="L"/>
  <fTestc="350" n="総コレステロール" Unit="mg/dl" Upper="220" Lower="120" DataType="NM" CodeType="L"/>
  <fTestc="353" n="LDL コレステロール" Unit="mg/dl" Upper="160" Lower="50" DataType="NM" CodeType="L"/>
  <fTestc="354" n="HDL コレステロール" Unit="mg/dl" Upper="85" Lower="30" DataType="NM" CodeType="L"/>
  <fTestc="351" n="中性脂肪" Unit="mg/dl" Upper="170" Lower="30" DataType="NM" CodeType="L"/>
  <fTestc="360" n="GOT" Unit="IU/l" Upper="40" Lower="7" DataType="NM" CodeType="L"/>
  <fTestc="361" n="GPT" Unit="IU/l" Upper="35" Lower="2" DataType="NM" CodeType="L"/>
  <fTestc="370" n="総ビリルビン" Unit="mg/dl" Upper="1.2" Lower="0" DataType="NM" CodeType="L"/>
  <fTestc="371" n="直接ビリルビン" Unit="mg/dl" Upper="0.4" Lower="0" DataType="NM" CodeType="L"/>
  <fTestc="700" n="尿糖" DataType="CE" CodeType="L"/>
  <fTestc="701" n="尿タンパク" DataType="CE" CodeType="L"/>
</Defaults>

```

図 7: デフォルト情報の記述例

## 4.2. 本体

本体“Body”では、健診データ“Report”が受診者数だけ、あるいは処理要求“Request”が必要数だけ繰り返す。なお、“Request”の下位構造は定義されていない。

```
<!ELEMENT Body (Request+|Report+)>
```

健診データ部分“Report”には受診者の個人情報である受診者データ“Patient”，検査データ“Test”，病歴データ“History”，問診データ“Query”，判定“Diagnosis”および医師の処方“Medication”などが含まれている。

```
<!ELEMENT Report (Patient,Test?,History?,
Query?,Impression?,Diagnosis?,Medication?)>
```

検査データ“rTest”について説明する。検査データの本質的な要素は計測データ値であり、他はその属性である。データの記述に際して、検査項目の出現順序や項目の種類を規定するとデータ構造を拘束することになる。そこで検査のデータ構造としては、“Test”配下に検査データのタグ“rTest”が任意回出現する単純なものとし、検査名などは属性として扱った。属性の構造は“fTest”と同じである。以下は検査データの記述例である。

```
<rTest c="400" n="血糖" Unit="mg/dl"
Upper="110" Lower="70" DataType="NM"
```

```
CodeType="JAHIS">100
</rTest>
```

もし、検査データの属性がヘッダ中のデフォルト情報と同じであれば、

```
<rTest c="400">100</rTest>
```

のように簡略化できる。HDMLでは他の規約のデータフォーマットに従った値も記述できる。例えばデータフォーマットが“ASTM4”であった場合は、

```
<rTest c="84925" n="Na" CodeType="AS4"
DataFormat="ASTM4">
OBX|100|ST|84295~Na||150|mmol/l
|136-148|H||A|F|
</rTest>
```

と記述できる。なお

```
CodeType="AS4"
```

はデータ項目を示すコード系が ASTM4 に基づいていることを示している。これをデータフォーマット“JAHIS”で記述すると単純に

```
<rTest c="411">150</rTest>
```

と記述される。

ところで、前述のように、データ項目には目的あるいは意味がある。合同委員会では健診データを機能とサブシステムという 2次元マ

|        |     |       |     |
|--------|-----|-------|-----|
| 日常生活動作 | 移動  | 寝返り   | 全介助 |
|        |     | 立ち上がり | 不可  |
|        |     | 歩行    | 不可  |
|        | その他 | 食事動作  | 全介助 |
|        |     | 整容動作  | 全介助 |
|        |     | 入浴動作  | 全介助 |

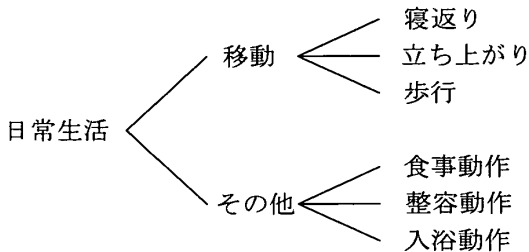


図 8: データの意味付けと構造化

トリクスでモデル化しているが、これはデータに対する意味付けに他ならない。データの目的や意味はマトリクスや表として表現することが多いが、図 8 に示すように、これらも樹形図として表現することができる。以上の点を考慮して、例えば問診 “Query” の DTD は以下のように定義されている。

```

<!ELEMENT Query((rQuery)*,(h0)*)*>
<!ELEMENT h0((rTest|rHistory|rQuery)*,(h1)*)*>
<!ELEMENT h1((rTest|rHistory|rQuery)*,(h2)*)*>
<!ELEMENT h2((rTest|rHistory|rQuery)*,(h3)*)*>
<!ELEMENT h3((rTest|rHistory|rQuery)*,(h4)*)*>
<!ELEMENT h4((rTest|rHistory|rQuery)*,(h5)*)*>
<!ELEMENT h5(rTest|rHistory|rQuery)*>
<!ELEMENT rQuery(#PCDATA)>
<ATTLIST h0 %atCommon; 1 CDATA #REQUIRED>
<ATTLIST h1 %atCommon; 1 CDATA #REQUIRED>
<ATTLIST h2 %atCommon; 1 CDATA #REQUIRED>
<ATTLIST h3 %atCommon; 1 CDATA #REQUIRED>
<ATTLIST h4 %atCommon; 1 CDATA #REQUIRED>
<ATTLIST h5 %atCommon; 1 CDATA #REQUIRED>
    
```

ここで要素 “h0” から “h5” は表の入れ子の深さを表し、これらの組み合わせにより複雑な表も記述することができる。以下の例では “h0” が表の最も左側の「日常生活」のレベルを表し、“h1” がその下位構造である「移動」と「その他」を表す。“h1” はそれぞれ「移動」と「その他」の下位構造を表し、データ値は “rQuery” の部分に記述する。なお、属性 “l (エル)” は

表の要素名を表す。

```

<Query>
  <h0 l="日常生活">
    <h1 l="移動">
      <rQuery c="RM1" n="寝返り">全介助
    </rQuery>
      <rQuery c="RM2" n="立ち上がり">不可
    </rQuery>
      <rQuery c="RM3" n="歩行">不可
    </rQuery>
    </h1>
    <h1 l="その他">
      <rQuery c="RE1" n="食事動作">全介助
    </rQuery>
      <rQuery c="RE2" n="整容動作">全介助
    </rQuery>
      <rQuery c="RE3" n="入浴動作">全介助
    </rQuery>
    </h1>
  </h0>
</Query>
    
```

### 4.3. トレイラ

トレイラ “Trailer” は交換されたデータをシステムが処理する上で必要な付加的情報を記載するものであるが、現在は未定義である。

## 5. HDML ツール

HDML の有効性を検証するためのデータ交換ツール（以下では HDML 変換ツール）を作成した [18]。ただし、HDML 変換ツールは評価用であるため、JAHIS 健診データ交換規約 Ver.1 の DTD で定義されている全ての要件を満たしてはいない。HDML 交換ツールはデータの送信側である健診施設とデータの受信側である健診依頼元の 2 カ所で利用される。データの流れを図 9 に示す。

データ送信側では、健診施設の健診情報システムからデータを抽出して CSV (Comma Separated Value) フォーマットのファイルを作成する。CSV ファイルには交換したいデータ項目のみを含めればよい。次に施設情報登録

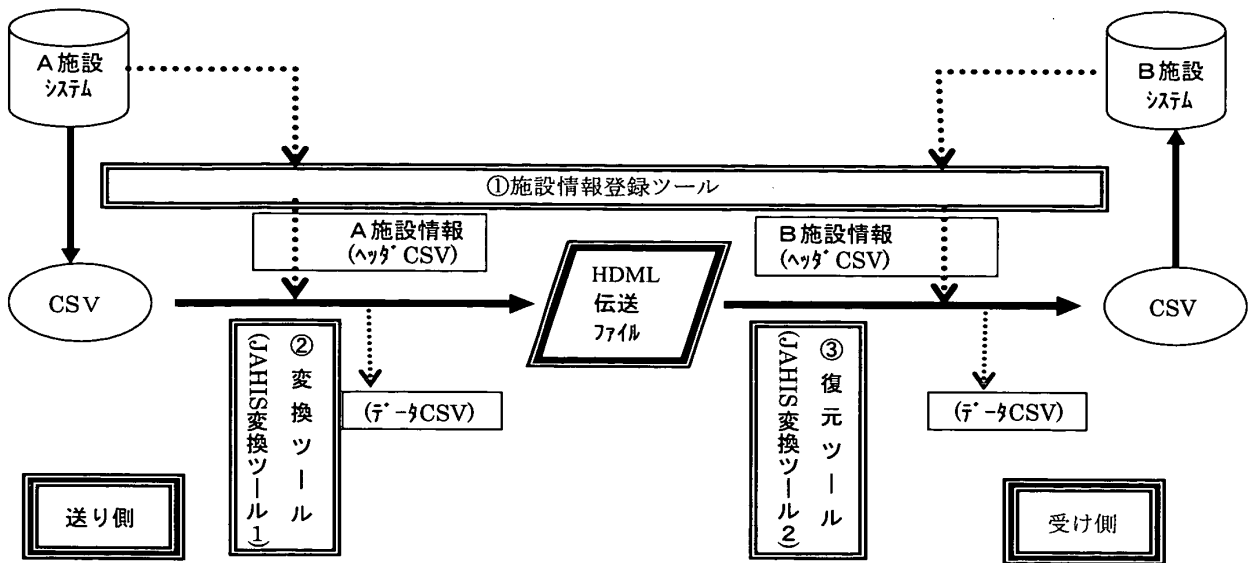


図 9: HDML 交換ツールの流れの概略図

ツールの支援を受けて属性情報定義ファイルを作成する。属性情報定義ファイルには、

- ① 健診機関の名称・住所など
- ② CSV ファイルのフォーマット定義
- ③ 各データの属性 (単位, 測定法など)

が含まれる。最後に CSV ファイルと属性情報定義ファイルを CSV-HDML 変換ツールに適用して HDML ファイルを生成する。

データ受信側では、施設情報登録ツールの支援を受けて、予め依頼元用の属性情報定義ファイルを作成しておく。送られてきた HDML ファイルとこの属性情報定義ファイルを HDML-CSV 変換ツールに適用し、依頼元の健診情報システムに適合した CSV フォーマットのファイルを生成する。依頼元の健診情報システムへのデータの取り込みは、この CSV ファイルを利用する。HDML 変換ツールの画面例を図 10 と 11 に示す。

## 6. 評価試験

HDML の有効性を検証するために、HDML 変換ツールを用いた評価試験を行った。

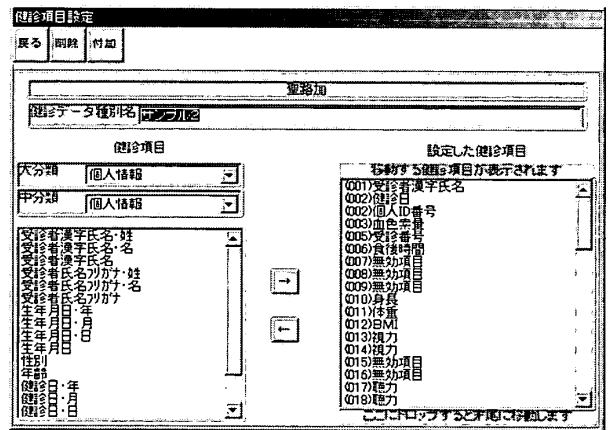


図 10: 健診項目設定ウィンドウ

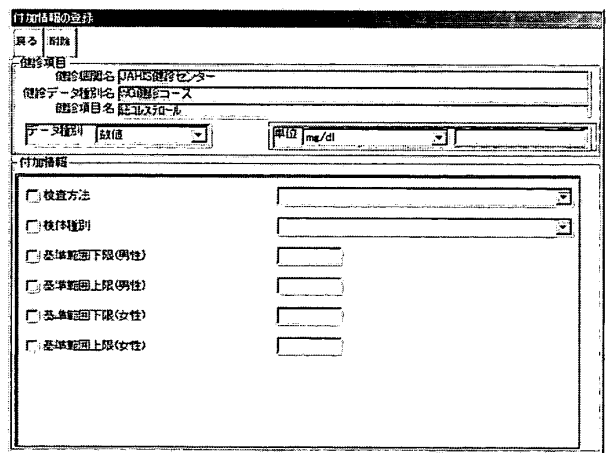


図 11: 付加情報の登録ウィンドウ

## 6.1. 実施場所

実施場所として、企業の健康管理センター（以下では施設 A）と人間ドック施設（以下では施設 B）の 2 カ所を選んだ。施設 A は企業内の健康管理センターで、社員の健診データを管理している。健診データは健診を委託した企業外の健診施設より提供される。施設 B は企業従業員や地域住民を対象とした人間ドックを行っている医療施設で、健診データは自施設のシステムで管理しており、依頼元企業に対しては健診データも提供している。なお HDML 変換ツールの操作は、施設 A では看護師が、施設 B ではシステムエンジニアが担当した。

## 6.2. 試験方法

以下の方法により、データ交換における再現性、互換性および操作性を検証した。

再現性の検証：施設 A と施設 B のそれぞれの施設で作成した健診データを、HDML 変換ツールにより CSV データから HDML データに変換する。次に同じ施設内で再び HDML から CSV に変換し、各施設の健診情報システムに取り込む。

互換性の検証：施設 A と施設 B のそれぞれの施設で作成した健診データを、HDML 変換ツールにより CSV データから HDML データに変換する。次にこれらのデータを施設間で交換し、HDML 変換ツールを用いて HDML データから CSV データに変換し、各施設の健診情報システムに取り込む。

## 6.3. 評価試験の結果

### (1) 再現性の検証

施設 A 内の交換試験では、251 件のデータ中 118 件について HDML データへ変換できた。変換に成功したデータについては、再び HDML データから CSV データへ変換した後、施設 A の健診情報システムに取り込んだ。その結果、118 件のデータについては内容に変化

を認めなかった。

施設 B 内については、208 件のデータ中 115 件について HDML データへ変換できた。変換に成功したデータについては、再び HDML データから CSV データへ変換した後、施設 B の健診情報システムに取り込んだ。その結果、115 件のデータについては内容に変化を認めなかった。

### (2) 互換性の検証

施設 A から施設 B へのデータ交換試験では、A 施設における再現性試験で成功した健診データ 118 件を利用した。このうち施設 B と交換して、直ちに施設 B の健診情報システムで使用できたデータは 48 件であった。

施設 B から施設 A への交換試験では、交換された健診データ 115 件のうち、48 件については交換して直ちに施設 A の健診情報システムで使用できた。

### (3) 操作性の検証

施設 A ではコンピュータ操作に習熟していない看護師が担当した。そのため、操作にかなりの時間を要した（試験後のアンケート調査より）。

施設 B ではコンピュータ操作に習熟したシステムエンジニアが担当したため、操作に関する問題は特に認められなかった。ただし語彙の設定では問題があった。つまり、JAHIS で定義した標準語彙と自施設の語彙をうまく関連付けることができず、一部のデータ項目を誤ったデータ項目に割り当てするなどしていた。これはシステムエンジニアが十分な医療知識を持ち合わせていなかったためと推測される（試験後の検討会報告書より）。

## 6.4. 評価試験の検討

### (1) HDML 化できなかったデータ

施設 A および施設 B の健診データのうち HDML 化できなかったデータを表 1 に示す。これらのデータには、HDML 変換ツールの設

表 1: HDML 化できなかったデータ種類

| データの種類 | 施設 A | 施設 B |
|--------|------|------|
| 問診データ  | 42   | 0    |
| 検査データ  | 67   | 66   |
| 判定データ  | 21   | 14   |
| 所属データ  | 3    | 13   |
| 計      | 133  | 93   |

計段階で除外されていた語彙（データ項目名や、問診・判定データ中に表れる用語および記号など）、あるいは変換ツールの語彙対応辞書が未整備で標準語彙に変換できない語彙があったため、HDML への変換に失敗したものであった。根本的な解決法は、語彙辞書の充実を図ることである。しかし問診データや判定データのように、多様な記述例あるいは記述法が存在する部分については更なる検討が必要であろう。

## (2) データ種別、表現方法の違い

HDML データには変換できたものの、交換したデータの半数以上が受信側で使用できなかった。これは受信側でデータの解釈に失敗したためである。例えば、送信側と受信側で同じデータ項目を扱っているながら、一方は数値データ、他方は定性データであるような場合で、現状の HDML では対応することができない。

データの表現方法の違いに起因していた失敗例もある。これは対象部位と所見を一体の情報として扱っている施設と、個別に扱っている施設があるような場合である。施設により「尿沈さ」の扱い方が異なっている例を表 2 と 3 に示す。

施設 A におけるデータ項目「尿沈さ計測数 1」は HDML では「尿沈さ」として設定できるが、「尿沈さ物質 1」と「尿沈さ視野 1」については語彙が用意されておらず、データ項目として書き込む場所がなく HDML への変換ができなかった。ただし JAHIS 健診データ交換規約 Ver.1 では、値の属性“Condition”

表 2: 表現の相違 (尿沈さ) 施設 A

|          |      |
|----------|------|
| 尿沈さ物質 1  | 硝子円柱 |
| 尿沈さ視野 1  | /W   |
| 尿沈さ計測数 1 | 1-4  |

表 3: 表現の相違 (尿沈さ) 施設 B

|      |      |
|------|------|
| 赤血球  | 1 以下 |
| 白血球  |      |
| 硝子円柱 | 1 以下 |

あるいは“Method”にこれらの情報を書き込むことができる。したがって、問題は HDML 変換ツールが HDML 規約を完全にサポートしていないためと言える。しかし HDML に従ってデータが記述できてデータ交換が成功したとしても、施設 A と施設 B のように相互の表記が全く異なっている場合、HDML データに変換はできても使用することはできない。これは今後の検討課題である。

施設 B では尿沈さ物質が決められているので、データ項目については設定が可能である。ただし、HDML 変換ツールは 4.2 で示したようにデータ項目を部位別に記述できないので、尿沈さの「赤血球」であるか血液中の「赤血球」であるかを区別することができなかった。これも HDML 変換ツールの問題である。

## (3) 担当者の習熟度および操作マニュアル

施設 A と施設 B で操作担当者を異なる職種とした目的は、HDML 変換ツールの操作性を検証するためであった。試験後のアンケート調査より、医療従事者を対象とするにはユーザーインターフェースの改良と簡便かつ詳しい操作マニュアルの作成が不可欠である事が分かった。

一方コンピュータ技術者を対象とした場合は語彙の理解に問題が認められた。この結果は、交換規約を使用する使用者の職種についても配慮する必要があることを示している。

## 7. 考察

評価試験により、JAHIS 健診データ交換規約 Ver.1 は再現性・互換性・操作性の全てにおいて解決すべき問題が存在するため、このまま現場へ導入することは難しいと思われる。しかし、これまでのデータ交換における膨大な時間と労力を考えれば、この評価試験により曲がりなりにも健診データの交換が実現できたことは、今後の施設間データ交換にとって大きな一歩であったと言える。

以下では HDML の課題と今後の展望についてまとめる。

### 7.1. 要素名について

HDML における SGML/XML 要素名は合同委員会において独自に決定されたものである。しかし、

- ① 名前が要素の内容を概ね表していること
- ② 半角英数文字を使うこと

以外には特段の命名規則は定められていない。また、ある要素名を構成する下位の要素名とその構造についても合同委員会で定義しているが、事例調査は網羅的ではなかった。その結果、HDML で定義された健康保険証に関する DTD 定義では、政府管掌健康保険データ中の保健記号、保健符号および保健番号のデータ項目を、明確に区別することができない。

このような問題の軽減を図るには、対応可能な要素名とその下位構造について、他規約との協調を図る必要がある。具体的には、「平成 11 年度 電子保存された診療録情報交換のためのデータ項目セットの作成 報告書」<sup>[19]</sup> および、ISO/IEC 11179 の “Specification and Standardization of Data Elements”<sup>[20]</sup> を参考として、HDML の要素名およびいくつかの下位構造の再定義を行う必要がある。

#### (1) 電子保存された診療録情報交換のためのデータ項目セットの作成報告書

医療機関相互の診療情報提供や行政機関へ

の届け出を電子的な方法で実施する場合を想定し、交換される情報を構成するデータ項目を送受信側とで共通化することを目的として作成された報告書である。この報告書で定義されている要素名は、既存の医療文書、医療情報交換用の規格、カルテの記載用紙などで使用されているデータ項目を多数収集し、整理・取捨選択されたものである。

報告書で定義されている要素名は、身体所見といった見出し語のレベルにとどまる。つまり「腹部圧痛」あるいは「胸部打診」といった身体所見に記載される所見項目は、「身体所見」に格納される値と考え、要素名としては定義されていない。そのため実際の情報交換にあたっては、交換する要素名が既に定義されている場合はそれ採択し、不足している要素名は独自に追加する必要がある。しかも独自に追加した要素名やその定義を登録・参照する方法など機能は示されていない。つまり、報告書に従って健診データを記述しても、独自の要素名が多くなり、問題の解決にはなりにくいおそれがある。

#### (2) Specification and Standardization of Data Elements

情報システム間のデータ交換で用いられる要素名の識別と標準化に関する規約であり、「電子保存された診療録情報交換のためのデータ項目セットの作成」とは異なり医療情報に限定されたものではない。例えば EDI(Electronic Data Interchange) の展開は標準データ項目の欠如により進展が遅れていると言われているが、本規約はこのような問題に対応したものである。情報システム間でデータを共有するためには、データの送受診側でデータの意味・表記・識別に関する共通の理解が確立されていなければならない。データの意味を理解するためには、データは適切に表記され、利用者がこれを取得する手段が用意されていなければならない。そのための仕掛けをデータ項目レジストリと呼んでおり、本規約はそのための標準化についての方法を提供する。



### (3) HDML における要素名の標準化

HDML と他規約との互換性を考えるうえで、要素名の共通化を推進することは必須である。これはデータ項目名の対応付けだけではなく、互換性を確保する上で、共有すべきデータの分類体系とデータの粒度を揃えるためである。例えば、患者の氏名や連絡先などの情報として、どの様な見出しレベルの項目名と、どの程度の詳細な下位構造が必要であるかを検討する上で「平成11年度 電子保存された診療録情報交換のためのデータ項目セットの作成報告書」は有益であると考えている。ただし項目名を日本語化するか否かについては意見の分かれるところである。ところで同報告書には別表2として区分表およびコード表が定義されている。HDML においても、データ型がコード型などの場合には利用可能なコード系を別表として掲載している。データ共有の観点からは、可能な限り同じ区分やコード体系に従うか、他規約の区分やコード系を取り込む必要がある。これについては早急に行う予定である。

要素名の共通化を本格的に行うにはレジストリ機能が必須であり、これについては“Specification and Standardization of Data Elements”が参考になろう。さらに、データの国際的な共有までを視野に入れるのであれば、要素名は英数文字で記述されていることが望ましい。したがって協調が可能な範囲において、HDML の要素名は ISO/IEC 11179 に適合させるべきであると考えられる。

## 7.2. 標準規約への対応

HDML は合同委員会で策定された健診データ交換規約である。HDML は、カードメディアに蓄積された健康データを交換するために研究されていた規約を、合同委員会の検討内容に合わせて拡張したものである<sup>[21][22][23]</sup>。一方、合同委員会の活動とは別に、JAHIS の臨床検査システム委員会では、HL7 をベースとした臨床検査データ交換規約の策定している

[24]。また医療情報学会を中心に MML(Medical Markup Language)<sup>[25]</sup> や MERIT-9(Medical Records, Images, Texts -Information eXchange)<sup>[26]</sup> など電子カルテの規約策定作業も行われている。合同委員会としても HDML を実装する上で、これらの動向を考慮する必要がある。以下では MML, MERIT-9 および HL7 について比較を行った。

### (1) MML/MERIT-9

MML は診療施設間で診療データを電子的に交換するための SGML/XML 形式のデータ記述言語である。メッセージの内容は、記録者の情報、患者の識別情報、生活歴、保険情報、既往疾患リスト、主訴、既往歴、家族歴、現病歴、検査結果、処方オーダー、処置記録、処置オーダー、サマリー情報など多岐にわたる。カルテが基本となっているため、一つのメッセージは一人の患者についての一つまたは複数時点での診療データを対象としている。

MML はテキストベースの電子カルテであるため、広範囲のデータを記述するためには多数のタグを必要とし、それらの維持管理は容易ではない。また画像などの非テキストデータの扱いも問題となる。そこで MML (Version 2.21) では XML Namespace を用いたモジュール化を進めている。

一方 MERIT-9 では、患者情報のフレームワークを定義し、検査結果や画像などのデータについては、HL7 あるいは DICOM (Digital Image and Communications in Medicine) など既存の規格によるデータファイルを外部参照する方式を採用している。

これらは日本における本格的な電子カルテ用の規約であり、今後の展開を注目する必要がある。しかし MML および MERIT-9 のいずれにおいても、検査結果などは既存の規約に基づいて作成されたデータを参照することになっている。つまり対象領域を健診に限定している HDML は、MML/MERIT-9 と競合する立場にはない。むしろ将来的には、HDML が電子カルテ規約のモジュールとして利用される可能性もある。

## (2) HL7

HL7は、患者の入院・登録、退院又は転院のデータ、照会、オーダー、結果、診療所見、請求書発行およびマスタファイル更新において、データを送受信する情報システム間のインタフェースを対象とした規約である。主要な目標は、医療コンピュータアプリケーションにおけるデータ交換のための規格を提供することである。HL7の利点は事実上のデファクト・スタンダードの一つである、対象が広範囲である、多くの標準化団体との協調を図っている点にある。

当初、HL7は合同委員会の要件を満たすものと考えられていたが、事例的検討を進める過程で機能の拡張が必要になった。このような場合、HL7のシンタクスを借用して独自の機能拡張を行い、可能であればHL7規約へ反映させる方法が考えられる。しかし実装がHL7のシンタクスと規約に拘束されて、拡張には不便であった。例えば健診データの場合、個人単位のデータ交換ではなく集団データの一括交換が主流である。HL7は一患者ごとのデータ交換が基本であり、健診データの交換には都合が悪い。またHL7は規模の大きな国際的な規約であるため、合同委員会の要求が、どれだけ速やかにHL7に反映できるかが問題となる。

HDMLは独自規約なので検討の結果を自由に反映させ、シミュレーションなどを容易に行うことができる。またデータ記述能力はHL7やMML/MERIT-9などと基本的には同レベルであり、HDMLデータの中核的な部分（例えばHL7のOBXセグメントと共通の部分）を、これらの規約に従ったデータに変換することは容易である。したがって当面はHDMLを推進し、最終的な判断はMML/MERIT-9やHL7などの動向を見ながら決定すればよいと考える。

## 7.3. 展望

標準健診データ交換規約の目的は、分散し

ている健診データを効率的に収集し、いわゆる生涯健康情報管理データベースというべきものを構築する基盤を提供することにある。これを基にして、多様なデータ解析による個人・集団の健康状態の評価・予測、さらに現在の疫学解析を超えたデータマイニングなどの手法による知見の獲得を行い、最終的には生涯健康管理の観点からライフサイクルにあった適切な健康指導を実現する。

そのためには正確なデータ交換を実現しなければならない。しかし6.4で述べたように、現状のHDMLは血清コレステロールのような数値データの交換には有効であったが、尿沈さのように表現法に多様性のあるデータについては対応できなかった。この問題を解決するには、データ項目名・対象物質・材料・部位などの関係を表現するモデルが必要である。これについてはLOINC (Laboratory Observations Identifier Names and Codes)<sup>[27]</sup>あるいは日本臨床検査医学会臨床検査項目分類コード(JLAC10)<sup>[28]</sup>の適用が考えられ、直ちに検討を始める必要がある。

HDML変換ツールでは、交換された数値データを変換するためにSDI (Standard Deviation Index)<sup>[29]</sup>を用いた。これにより単位や計測法などの相違に由来する数値データのズレを修正することができた。これは

$$SDI = (Data - Middle\ of\ Range) / Range \times 40 + 100$$

で計算される。なおRangeは検査値の正常範囲の上限値と下限値の差である。SDIにより、全ての検査値は、もし値が正常範囲の下限であれば80、上限であれば120に標準化される。なお検査値の正常範囲は、図7に示すように、HDMLデータ属性の“Upper”と“Lower”に記載されている。これ以外にも

$$common\ unit = (Data - Median) \times 10 / SD + 100$$

で計算する方法などがある<sup>[30]</sup>。なお上式においてMedianは標準Median、SDは標準偏

差を表す。どのような変換法が適当であるかについては評価を行う必要がある。また将来的には個人ごとのデータ分布に基づいた評価を行うべきであると考えている<sup>[31]</sup>。

HDMLの構造定義には技術的に未熟あるいはアドホックな部分があり、これがデータ処理を複雑にしているなどの問題点も明らかになった。そこで前述の要素名の標準への対応を含めてHDMLのDTDを修正しつつある。例えば4.2(HDMLの本体部分の説明)で述べた階層的な表の定義では、階層の深さを先験的に最大で6などとしていたが、これを

```
<!ELEMENT Query((rQuery)*,(h)*)*>
<!ELEMENT h((rTest|rHistory|rQuery)*,
(h)*)*>
```

のように変更し、可読性の改善と適用範囲の拡大を図った。また4.1(HDMLのヘッダ部分の説明)で述べたように、HDMLでは、各データに共通な属性情報をヘッダの“Defaults”部分に、異なる属性情報をデータ本体の“Body”部分に別々に記述していたが、これがデータ処理を複雑にしていた。そこでデータに関する全属性情報をヘッダ部分に集めることにより、データ処理の効率化とデータ本体の可読性の向上を図った。

## 8. 結語

日本総合健診医学会情報委員会は、健診データの有効利用を図るため、日本保健福祉医療情報システム工業会と合同委員会を組織し、標準健診データ交換規約HDMLとツールの開発に着手した。HDMLはSGML/XMLを基礎とした健診に特化したデータ交換規約である。HDMLの導入により、データ変換に伴うコストの削減が期待される。

健診情報システム間においてHDMLを介したデータ交換を行うためにHDML変換ツールを開発し評価試験を行った。その結果、数値データに限定すれば健診情報システム間におけるデータ交換は成功し、HDMLは有効であることが確認された。しかし健診データ全

般については解決すべき問題が存在し、このまま現場に導入することは難しいことも明らかになり、今後の課題となっている。

このように評価は限定されたものであったが、これまでのデータ交換における膨大な時間と労力を考えれば、曲がりなりにも健診データ交換が実現できたことは大きな一歩であったと言える。

## 参考文献

- [1] 日本保健医療情報システム工業会(JAHIS): <http://www.jahis.or.jp>.
- [2] 日本保健医療情報システム工業会:「平成8年度健診データ標準化活動報告書」, 1997.
- [3] Hara, S; H, Sugomori; K, Yoshida; I, Tofukuji; T, Kubodera: “Health Data Transfer Protocol based on SGML - The Health Data Markup Language (HDML)-”, IEEE EMBS International Conference on Information Technology Applications in Biomedicine (ITAB-IT 2000) Proceedings, pp.256-261, 2000.
- [4] S, Sugimori; K, Yoshida; S, Hara; K, Furumi; T, Kubodera; I, Tofukuji: “Standard Protocol for Exchange of Health-checkups based on SGML: The Health-checkups Data Markup Language (HDML)”, Proceedings of An Information Technology in Community Health 2000 (ICTC2000), 2000.
- [5] The International Organization of Standardization (ISO): “ISO 8879: Information processing - Text and office systems - Standard Generalized Markup Language (SGML)”, 1986.
- [6] 日本規格協会:「JIS X 4151-1992 文書記述言語 SGML」, 1992.
- [7] World Wide Web Consortium (W3C): <http://www.w3c.org/TR/REC-xml>.
- [8] Health Level Seven: “HL7 Version 2.3”, 1996.

- [9] American Society for Testing and Materials (ASTM): "Standard Specification for Transferring Clinical Laboratory Data Message Between Independent Computer Systems ASTM E1238-88", 1994.
- [10] 財団法人医療情報システム開発センター: 「臨床検査データ交換規約 (暫定版)」, 1993.
- [11] Georges J.E. De Moor; Clement J. McDonald; Jaap Noothoven van Goor: "Progress in Standardization in Health Care Informatics", IOS Press, 1993.
- [12] 杉森裕樹; 吉田勝美; JSHIS 合同委員会: 「健診情報委員会 - 健診データ伝送規約に基づく健診データ変換システム (Health-checkup Data Markup Language: HDML) -」, 日本総合健診医学会誌, Vol.27, No.3, pp.309-316, 2000.
- [13] Dudeck, Joachim: "Aspects of implementing and harmonizing health-care communication standards", International Journal of Medical Informatics, No.48, pp.163-171, 1998.
- [14] 厚生省健康政策局総務課医療技術開発室 (監修): 「保健医療カードシステム標準化マニュアル第一版」, 財団法人 医療情報システム開発センター, 1994.
- [15] World Health Organization (WHO): "International Statistical Classification of Diseases and Related Health Problems, 10<sup>th</sup> Revision", 1992.
- [16] American College of Pathology: "Systematized Nomenclature of Medicine (SNOMED) International. Version 3.3", 1996.
- [17] 渡辺; 木村; シルチン; 谷: 「ICD-10 + 診療報酬請求病名集と SNOMED International とのマッピングと構造化提案」, 第 17 回医療情報連合大会論文集, 2-G-2-3, 1997.
- [18] 吉田勝美; 田村政紀; 菅原源二; 稲田紘; 原正一郎; 古海勝彦; 杉森裕樹; 東福寺幾夫; 窪寺健; 与田尊司; 高橋為生: 「JAHIS 健診データ交換規約について - 健診データの施設間共有の標準化について -」, 労働衛生管理, Vol.12, No.4, 2001, pp.18-24.
- [19] 財団法人医療情報システム開発センター: 「平成 11 年度電子保存された診療録情報交換のためのデータ項目セットの作成報告書」, 2000.
- [20] The International Organization of Standardization (ISO): "ISO/IEC 11179-1 Information technology - Specification and standardization of data element Part1: Framework for the specification and standardization of data elements", 1999.
- [21] Hara, Shoichiro; *et. al.*: "An Application of Optical Cards to Mass Health Examination", Proc. 6<sup>th</sup> Conf. on Med. Info., pp.1164-1168, 1989.
- [22] Hara, Shoichiro; *et. al.*: "Advantages of Optical Cards in Health Care"; Ann. Inter. Conf. of the IEEE Eng. Med. and Biol. Soc., Vol.13, No.3, pp. 1389-1390, 1991.
- [23] Hara, Shoichiro: "Health Care Support System Based on Optical Cards", health Cards '95 in Technology and Informatics, pp.294-298, IOS Press, 1995.
- [24] 日本保健医療情報システム工業会: 「平成 8 年度臨床検査データ交換規約 JAHIS-DRAFT」, 1997.
- [25] 吉原博幸; 大江和彦; 大橋克洋; 山本隆一; 山崎俊司; 廣瀬康行; 松井くにお; 日紫喜光良; 山下芳範; 皆川和史; 小山博史: 「医療情報の交換手順の標準化に関する研究報告書」, 1997.
- [26] 木村道男; 大江和彦; 吉原博幸: 「医療情報交換規格運用指針 MERIT-9 の概要」, 第 17 回医療情報連合大会論文集, pp.564-565, 1997.
- [27] LOINC Committee: "Logical Observations Identifier Names and Codes

(LOINC) Users Guide”, 1995.

[28] 日本臨床検査医学会：「臨床検査項目分類コード 第10回改訂」, 1997.

[29] 丹波正治, 清水裕史, 高橋為生：「臨床データの共通表現法」, 臨床病理 (臨時増刊), No.84, pp.39-59, 1990.

[30] G. Suganuma: “Proposal of the common unit for forecasting of health trend”, Japanese Journal of MHTS, No.26, pp.250-282, 2000.

[31] 上野春樹；根本幾；斉藤正男；岩塚徹：「AMHTSデータの信頼性—個人分布に基づく評価の検討」, 日健診誌, Vol.3, No.2, pp.15-22, 1976.

(2002年10月24日受付)

(2002年12月13日採録)

# 「XMLに基づくコンピュータ処理可能な構造化デジタル オブジェクト」の概念による Web リソースの 検索システムの研究と開発

## Research and Development of a Web Retrieval System on the Concept of “Computer-Processable Structured Digital Object based on XML”

李 穎\* 石塚英弘\*

Ying LI and Hidehiro ISHIZUKA

膨大な学術情報を如何に表現し、提供し、また如何なる検索機能を提供すれば良いか。これは古くかつ新しい課題である。これまで種々の手法が開発されてきたが、それぞれに欠点があり、研究者、すなわち、特定の情報源に限定されることなく、広く情報を調査し、よりノイズの少ない状態で有用な情報を得て、得られた結果を自らの観点から再構成しようとする利用者にとって、満足できる状況とは言いがたい。これを解決するためには、新たな概念に基づく方法論が必要であると考え。本研究では、「XMLに基づくコンピュータ処理可能な構造化デジタルオブジェクト」、すなわち、情報の最小単位を構成要素とし、XMLを用いて表現する構造化情報表現を考え、その概念に基づいて、Web リソースの検索システムのモデルを考案した。そして、典型的なファクトデータベースである Merck Index のデータを例として、プロトタイプを開発した。そのデータ構造は XML schema を用いて表現し、ASP.NET 上に実現した。

How is huge scientific information described and serviced? What retrieval functions should be provided? These are old and new problems to be solved. So far various methods have been developed, however each of them includes some defects. The situation is not satisfying for the researcher's needs who wants to investigate information widely not restricted to the particular resource, to get valuable information with less noise, and to reorganize obtained information according to his/her view. A new methodology on new concept is necessary to resolve the situation. We proposed concept of “Computer-Processable Structured Digital Object based on XML (Extensible Markup Language): (CPSDO/XML)” that is composed of elementary units of information, and designed an information retrieval system based on CPSDO/XML for Web resources. We also developed a prototype for Merck Index data, that is a typical fact database, on ASP.NET. The data structure is described with XML schema.

キーワード: Web 情報検索システムモデル, 検索対象, デジタルオブジェクト, XML, XML schema, XSLT, XPath, DOM, ASP.NET, Web information retrieval system model; retrieval target; digital object

### 1. はじめに

膨大な学術情報を如何に表現し、提供し、また如何なる検索機能を提供すれば良いか。これは古くかつ新しい課題である。これまで

種々の手法が開発されてきたが、それぞれに欠点があり、現時点では全ての問題点を解決できるシステムは見当たらない。そのため、さらに良い方法を求めて研究開発がなされている。我々もこのテーマを研究しており、本論文でその成果を報告する。

\* 図書館情報大学

University of Library and Information Science

まず、これまで行われてきた方法、および最近提案された方法のうち代表的なものを以下に挙げ、それらの長所短所を述べる。

- (A) 文献情報データベースとその情報検索システム
- (B) ファクト・データベースとその検索システム
- (C) 引用関係を利用した文献検索システム
- (D) 人工知能のフレームによる情報知識表現
- (E) 多数のファクト・データベースと文献情報データベースを搭載した検索システム
- (F) Web-HTMLによる文書群と Web 検索エンジン
- (G) オブジェクトデータベースによる表現と検索
- (H) オンライン・ジャーナル
- (I) XML(Extensible Markup Language)
- (J) Topic Map<sup>[1]</sup>
- (K) Semantic Web<sup>[2][3]</sup>

(A) の典型例は、化学分野では CAS (Chemical Abstract Service)<sup>[4]</sup>による抄録、事項索引、化学物質索引、書誌情報のデータベースとその情報検索システムであり、医学分野では米国国立医学図書館 (National Library of Medicine) による MEDLINE, Elsevier Science<sup>[5]</sup>による EMBASE (Excerpta Medica database) などである。これらは有効で広く使用されたが、これらの方法には限界があった。すなわち、使用料金の高さ、原文献の出版とデータベース公開とのタイムラグ、検索結果と文献が電子的に接続していなかったために検索結果を基に文献を書棚から探し出す必要があること等である。最後の点は全文データベースあるいはオンライン・ジャーナルとの連携がその後実現して解消し、タイムラグは出版段階からデータを文献情報データベース作製機関に送信することによる改善が見られるものの、事項索引など高度な索引の場合にタイムラグを無くすことは困難である。また、料金の高さの問題は現在も未解決である。

(B) は整理され、信頼できる情報を得ようとする利用者のために作られた便利なデータ

ベースであるが、使用料金の高さやタイムラグが欠点である。また、ファクト・データベースはそれぞれの対象分野・主題があり、収録しているデータ項目、内容は微妙に異なる。そのため、利用者は自分が得たい情報が収録されているファクト・データベースはどれかを知る必要があり、また複数のファクト・データベースから得られる結果を利用者が手作業で整理・組み合わせなければならない。

(C) の典型例は米国 ISI<sup>[6]</sup>による Science Citation Index であり、引用関係で検索できる点が特徴である。或る文献に引用された文献だけでなく、引用した文献も検索することができ、それを繰り返すことによって関連する文献群を得ることができる。ただし、引用索引データベースの編集はコストが掛かり、タイムラグも発生し、また利用料金もかなり高い。

(D) 人工知能の分野では情報知識の表現にフレームが用いられてきた。Minsky<sup>[7]</sup>が提唱したフレームは、その中にスロットを持つことによって、知識を適度な大きさのまとまりとして表現し、また、スロットに他のフレームへのリンクを付加することによって知識を構造化する。スロットには付加手続きを付けることもできる。フレームの実装としては、KRL<sup>[8]</sup>, FRL<sup>[9]</sup>等々様々なシステムが作られたが、そのことは逆に夫々のシステムで記述された知識/情報データの交換利用が難しいことを示した。また、百科事典の知識を記述しようとした Lenat の CYC プロジェクト<sup>[10]</sup>は大量の知識/情報を 1 箇所で作ることの難しさを浮彫にした。大量の場合は、分散して作製された知識/情報を互いに交換利用する方が実践的であるが、フレームではそれが困難なことが欠点である。

(E) の典型例は STN International<sup>[11]</sup>である。ファクト・データベースだけでなく、文献情報データベースも在る点が特徴である。個々のデータベースを検索でき、また統合検索機能を用いれば、複数のデータベースを同時に検索することができるが、得られる結果は人間が見て分かる形式であり、利用者側の

ソフトウェアによってデータ項目ごとに、あるいはそれらを組み合わせて二次処理することは容易でない。また、使用料金の高さは依然として解消されていない。

(F) は、全文が得られる点、関連ページに跳ぶリンク機能がある点が有用である。しかし、HTML で記述されているため、(A)、(B)、(E) のようにデータ項目が区別されていない。また、Web 検索エンジンは HTML を対象としているため、データ項目を区別することができない。その結果、検索ノイズが頻発する。これは明らかな欠点であり、解決されていない。

(G) オブジェクトデータベースは、オブジェクト指向データベースあるいは object-based database のことであり、プログラミング言語の観点とデータベースの観点の 2 つの面がある。後者では、オブジェクト間の関係によってデータ構造をモデル化する。知識を構造化して表現する点では (D) のフレームの影響を受けている。オブジェクトデータベースの例としては、O2<sup>[12]</sup> や ODMG 標準<sup>[13]</sup> などがある。複雑な構造を持つ情報の場合は複合オブジェクトを用いることになるが、学術情報のように複雑な場合は複合オブジェクトを用いても容易ではない。オブジェクトデータベースの場合も処理系の間でのデータの互換性の問題が存在する。

(H) は全文が得られる点、リンク機能を用いて参照文献のページに跳ぶことができる点が有用である。この 2 点は (F) と同じである。さらに、(A) の文献情報データベースと連携すれば、データベースの検索結果から当該ジャーナル論文のページに跳ぶことができ、この点は (F) よりも便利である。もっとも、実際にそのページに跳ぶためには、そのページの著作権を自社で持つか、あるいは、著作権を持つ出版社あるいは学会との協定を結ぶ必要がある。それは「寡占状況における情報の囲い込み」という別の大問題を引き起こした。たとえば、Elsevier Science 社は自社のみで大部分の情報を提供しようとする戦略を持ち、ヨーロッパの老舗の出版社を次々買収して傘下に

収めた。その結果、同社が提供するサービス：Science Direct<sup>[14]</sup> は圧倒的多数のオンライン・ジャーナルを以って様々な学問領域をカバーし、また、EMBASE も以前から持ち、前述の機能を実現している。これは、他のオンライン・ジャーナルを提供する出版社・学会にとっては脅威である。そのため、米国の大学会などは互いのページの参照を認め合う協定の締結を進めて対抗しているが、状況は厳しいようである。また、学術雑誌の価格の高騰が問題になっているが、その一因は寡占状況にあると推測されている。

(I) XML (Extensible Markup Language)<sup>[15]</sup> は、情報の階層構造、同じ階層の上の要素間の順序関係、関連情報へのリンクなどが表現できることが特長である。XML は単純なデータだけでなく、ドキュメント、テキストにも適用できる。XML による表現は処理系に依存しないため、データ交換に適している。また、Web を想定して設計された仕様であるため、URI を介して、インターネット上に分散して存在する情報を取得することができる。また、スキーマを自分で設計でき、適用分野や適用対象に制限は無いが、このことは逆に分野や対象に適合したスキーマを設計する必要があることを示している。

(J) は、トピックマップの作成者が認識している概念体系に則って、情報リソースを分類、整理、管理、検索、ナビゲートする一つのパラダイムである。情報リソースとは独立したメタレイア（上位層）として、個々のトピックを互いに関連付けた一種のトピック間関連図を構成し、個々のトピックから関連する情報リソースに参照のリンクを張る。トピックには、情報リソースが持つ主題を選んでもよいし、概念を選んでもよい。トピック間関連図により、主題と概念の間関係と概念間関係を示す。情報リソースの種類は問わず、ドキュメントであってもよいし、データベースであってもよい。また、情報リソースには個々のトピックから単にその該当個所へのリンクを張るのみであって、情報リソースを加工す



ることではない点の特徴である。しかし、このことは逆に得られる情報を加工・再構成する機能は無いことを示している。

(K)は、次世代の Web とされ、データの記述言語に HTML ではなく、データ項目を区別できる XML を用いること、データ項目間の関係を RDF (Resource Description Framework)<sup>[16]</sup>で記述すること、特定の Web サイトに集中するのではなく、分散型であり、各 Web サイトで異なる用語を用いることを許し、Ontology を用いることによって用語の問題を解決しようと考えていること、検索には Web 検索エンジンではなく、エージェントを用いることなど興味深いパラダイムである。しかし、現時点では構想段階であって、多くの未解決の先進技術に依存するため、実装までは至っていない。

(A) から (K) に述べた現状は、研究者、すなわち、特定の情報源に限定されることなく、広く情報を調査し、よりノイズの少ない状態で有用な情報を得て、得られた結果を自らの観点から再構成しようとする利用者、にとって満足できる状況とは言いがたい。そのためには、新たな概念に基づく方法論が必要であると考える。

そこで 2002 年 7 月、我々は情報の構成単位として CPDO (Computer Processable Digital Object) の概念を考案し、利用者が検索によって得たい情報の対象、すなわち「検索のターゲット」は CPDO によって表現できることを述べ、また、様々な分野の情報が CPDO を基に再構成できる可能性を示し、CPDO の概念に基づくデジタル図書館システムのモデルを提案した<sup>[17]</sup>。その後、我々は、CPDO の概念規定を改良して、CPSDO/XML (Computer-Processable Structured Digital Object based on XML) 「XML に基づくコンピュータ処理可能な構造化デジタルオブジェクト」を考案した。また、その概念に基づいて、Web リソースの、様々な情報リソースの典型例を採り上げて、その中のデータ項目の CPSDO/XML への変換、CPSDO/XML に依る検索システム

のプロトタイプの XML 処理系: ASP.NET<sup>[18]</sup>上での作成などを行い、前述の課題に対して一定の成果を得たので、本論文でそれを報告する。

## 2. CPDO および CPSDO/XML の概念またそれに基づく検索システムのモデル

本論文では、利用者が扱う学術情報の最小単位を CPDO (Computer Processable Digital Object) という。また、CPDO を構成要素とする構造化情報表現を CPSDO (Computer Processable Structured Digital Object) という。たとえば、「検索のターゲット」すなわち「利用者が検索によって得たい情報対象」が情報の最小単位であれば、それは CPDO であり、もしも、CPDO を組み合わせたものであれば、それは CPSDO の一部分である。

CPDO と CPSDO は XML によって記述されるデジタルオブジェクトであり、XML のタグ名称は記述内容を示す。XML で記述された CPSDO を CPSDO/XML (CPSDO based on XML) という。そのため、CPSDO/XML はソフトウェアによる処理が可能である。また、もしも検索対象の Web ページあるいはデータベースが CPSDO/XML に基づいて構築されるようになれば、CPSDO/XML の該当部分のみを検索するだけでよくなるため、ノイズの少ない効率的な情報検索が可能になる。

なお、CPDO および CPSDO で言うオブジェクトとは XML で言うオブジェクト (物あるいは対象とする構成要素) であって、メソッドは持っておらず、その点でオブジェクト指向で言うオブジェクトとは異なる。

我々の先の論文<sup>[17]</sup>では、CPDO と CPSDO を区別せず、CPDO の中に CPSDO を含めていた。また、CPDO の粒度は小さくても良いし、大きくても良いとしていた。利用者が検索して得たい情報の粒度は様々だからである。検索対象として論文の場合を検討し、大きい方から小さくなる順に、ある主題 (たとえば、

炭疽菌に効く薬)に関する論文の集合, 同じく1つの論文, 論文の一部(たとえば, 薬の薬効を記述した部分)の集合, 同じく論文の一部, 語(たとえば, 薬の名称)の集合, 語などから成る階層構造を考えたが, 対象となる文献は論文以外に, 単行本, 辞書, データベース, Web ページなどがあり, そのため, 様々な粒度がありえることになる. たとえば, 炭疽菌に効く薬の名前を検索したい, その薬の製法を検索したい, 製法が書いてある文献の書誌情報が欲しい, あるいは, 製法が書いてある記述そのものを得たい, あるいは製法のエッセンスである化学反応式を得たい, あるいはそれらが記述された文献を得たい, 等々の要望がありえる. これらに対応する CPDO の粒度は異なる. しかし, 論文<sup>[17]</sup>の考え方には分かり難い点があるため, 本論文では CPDO と CPSDO を区別することにした.

また, CPSDO を XML で記述することによって, CPSDO と CPDO の間に階層関係が存在する場合は, それを XML の階層関係を用いて表現できる. また, CPDO と CPSDO の間にリンク関係があれば, それを XML のリンク関係として表現することができる.

このように本論文では前論文の考え方をさらに進めて, 既存の情報リソースのデータ項目を XML 要素で記述した CPSDO/XML あるいはその構成要素である CPDO へ変換し, これにアクセスすることによって検索対象となった CPDO を取り出すことを可能にした. その結果, ①単一の CPDO による検索のみならず, ②複数の CPDO を組み合わせた検索, ③他のリソースの該当箇所への連鎖検索などが可能な検索システムのプロトタイプを作成した. 図1にそのシステムのモデルと機能を示す.

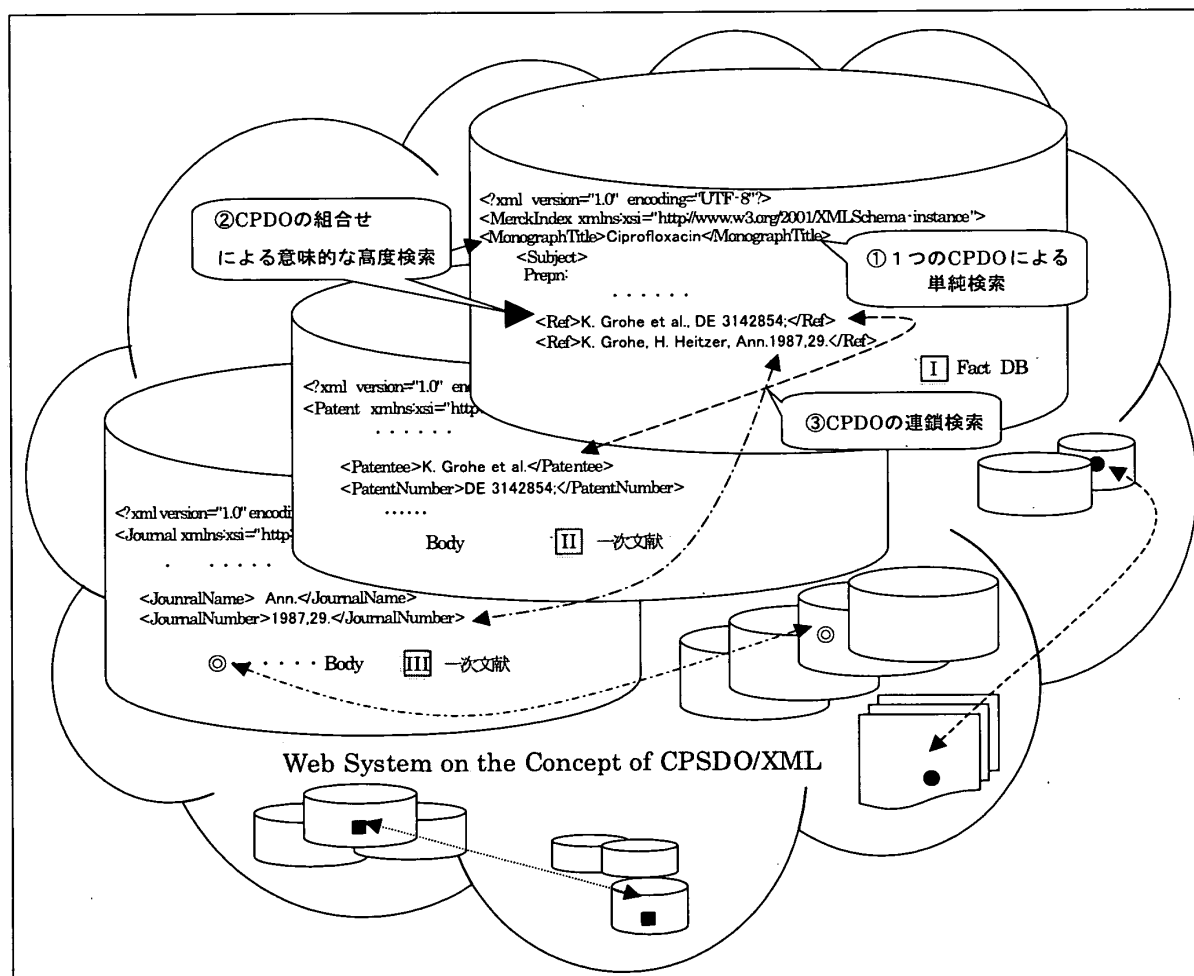


図1: CPDO および CPSDO/XML に基づく検索システムとその機能 (概念図)

図1のⅠFact DBは化学物質・医薬品の辞書またはファクトDBとして定評のあるMerck Index<sup>[19][20]</sup>の各データ項目を、対応するCPDOに変換し、XMLで表現したものである。これはWebサーバ上のXML処理系を用いることにより、検索することができる。実例は本論文の中で後に詳しく述べる。また、Ⅱ一次文献は特許のDBのCPSDO/XML版であり、Ⅲ一次文献は雑誌文献のDBのCPSDO/XML版である。なお、ⅡとⅢは説明用の仮のものであって、実際に変換作成したものではないが、変換作成は原理的に可能であり、検索も可能である。

#### ①1つのCPDOによる検索

例えば、抗生物質名“Ciprofloxacin”(炭疽菌に効く薬)の検索である。図1のⅠFact DBの該当するCPDOのXMLタグ: MonographTitleを指定して、その値が“Ciprofloxacin”であるものを検索すればよい。こうすることにより簡単にノイズレスで検索目標を達成できる。

#### ②CPDOの組合せによる意味的な高度検索:

例えば、「Ciprofloxacinの製法」を検索したいときは、図1のⅠFact DBのCPDOのXMLタグ: MonographTitleの値が“Ciprofloxacin”であるものについて、その下のレベルのCPDOがSubject, その下のCPDOがPrepn.(preparation:製法)の下のCPDOがRef(参考文献)のものをアクセスすることによって、製法が書かれた文献の書誌情報を知ることができる。このようにCPDOを組み合わせることによって「Ciprofloxacinの製法」のような意味検索を実現することができる。

#### ③CPDOの連鎖検索:

従来の情報検索システムでは、二次文献DBにより一次文献のIDを入手しても、両者は接続していないことが少なくなく、一次文献を得るには手間が掛かった。しかし、CPSDO/XMLに基づいた図1に示す検索システムが成立すれば、「Ciprofloxacinの製法」を記述した文献のXMLベースのDB、

すなわち、Ⅱ一次文献(特許)あるいはⅢ一次文献(雑誌)の該当文献へリンク機能を用いて直接行くことができる。これは連鎖検索と言えよう。また、一次文献中の参照文献リストのCPDOから他の一次文献の全文DBのCPSDO/XMLのCPDOにリンクすれば、検索の連鎖はさらにつながることになる。図1の◎と←→あるいは■と←→は、それを示している。

### 3. CPDO および CPSDO/XML に基づく検索システムの設計

図2の上半分に示すように、CPDOおよびCPSDO/XMLに基づく検索システムにおける「検索」とは、図2の④「CPSDO/XMLのXMLスキーマ(XML Schema<sup>[21]</sup>)」を用いて、XMLで記述された情報リソース、たとえば、Merck Indexから、⑤「検索ターゲットとなったCPDOを取り出すこと」である。この検索システムによって、ニーズの異なる様々な検索者にノイズの少ない検索結果を与えることが可能になる。

また、この検索システムを構築するには、まず、CPDOから構成される情報の構造(スキーマ)を設計し、次いで、各CPDOの値を既存の情報リソースの対応するデータ項目から得て、XML記述形式に変換すること(図2の下半分参照)が必要である。既存の様々な情報リソースのデータ項目は、CPDOのそれと一致するとは限らない。なぜなら、既存のリソースのデータ項目は人間(検索者)が見て理解できればよいため、ある程度まとまった単位の情報となっており、必ずしも情報の最小単位ではないからである。CPDOの考え方から言えば、2節で述べたように、CPDOは情報の最小単位であり、もしも、それより大きな単位が必要であれば、CPDOを組み合わせることで大きな単位のCPDOを作ることになる。この考え方に依り、CPDOから構成される情報のスキーマを設計する。

各CPDOの値を既存の情報リソースの対応

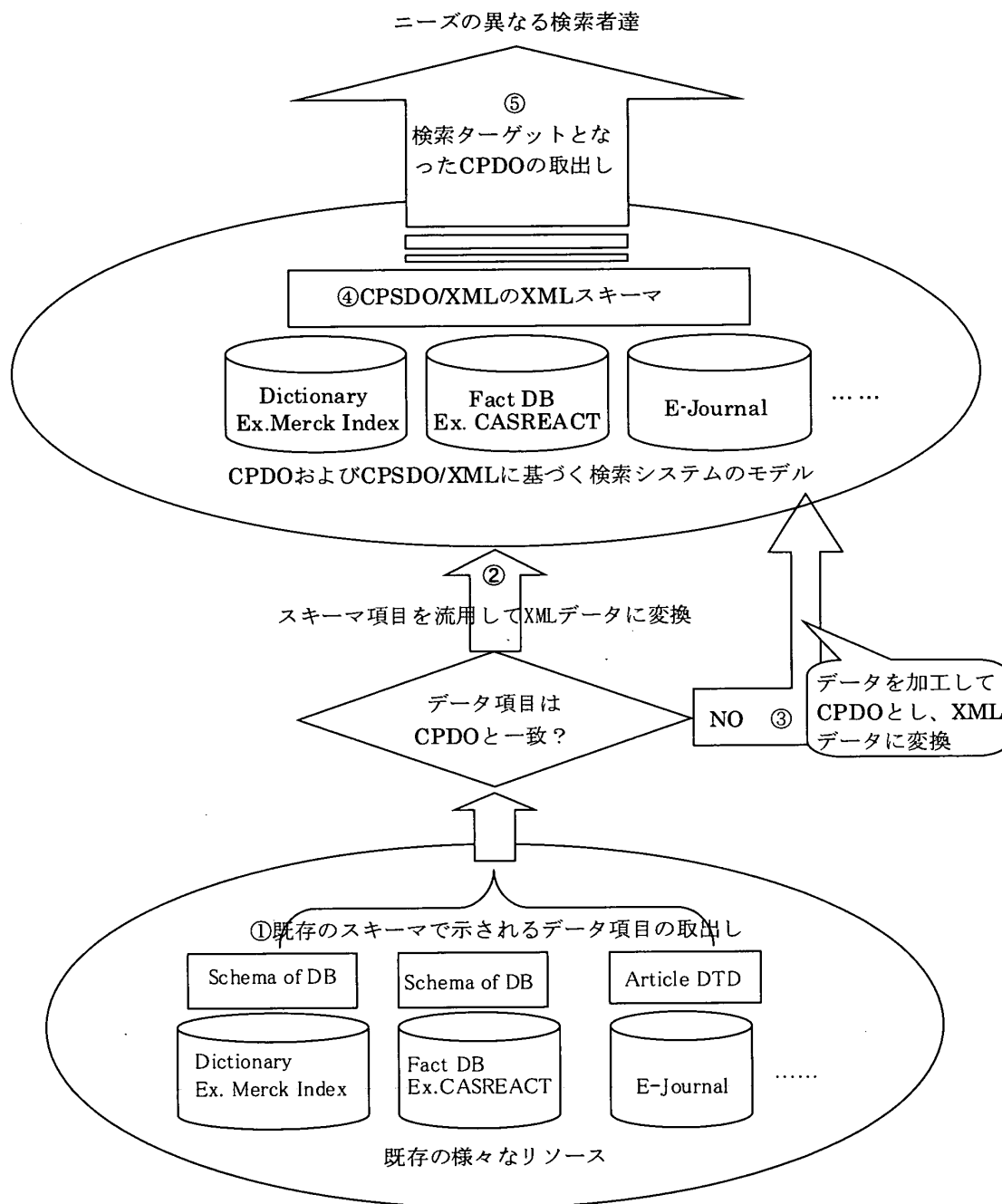


図 2: CPDO および CPSDO/XML に基づく検索システムとその構築の概念図

するデータ項目から得て、XML 記述形式に変換するには、次に示す処理を行う。なお、具体例は次節に述べる。

- ①既存のリソースから、そのスキーマに従ってデータ項目を取り出す。
- ②取り出したデータ項目が CPDO と一致するか否かを判定し、YES ならば、そのスキーマの構成要素単位を流用して対応する CPDO の XML データに変換し、XML ベースのリ

ソースに格納する。

- ③判定結果が NO ならば、データを加工して対応すべき CPDO に一致させ、XML データに変換して該当する XML ベースのリソースに格納する。

#### 4. CPDO および CPSDO/XML に基づく Web 検索システムの実装

我々は、既存のリソースの典型例として、薬学および化学の分野で最も広く利用されているリソースの一つである Merck Index<sup>[19]</sup> を採り上げ、前節で述べた設計方針による実装を行った。以下に述べる実装方法は勿論、他の種類の既存リソースにも適用可能である。

##### 4.1. 実装の概要

Merck Index のデータ項目を冊子体の記述を参考にしつつ、情報検索システム：DIALOG の Merck Index データベースのスキーマに依って分析した結果、DIALOG システムにおけるデータ項目の一部は、利用者が検索ターゲットとしたい情報の単位：CPDO に比べて大きいことが明らかになった。そこで、CPDO に比べて大きい単位のデータ項目は CPDO の単位に細分化し、CPDO と同じ単位のデータ項目はそのまま CPDO に対応させて、CPSDO/XML の XML スキーマを設計した。なお、本論文で設計した XML スキーマについては後で述べる。

次いで、既存のリソースのデータ項目の値を対応する CPDO の値に変換し、CPSDO/XML の XML DB を作成した。そして、この XML DB を持つ Web サーバを ASP.NET を用いて実装した。この検索システムは Web ブラウザ：Microsoft I.E. を用いて検索でき、検索者が必要とする CPDO を XML DB から取り出し、検索結果として表示する。

図 3 に前述のプロセスを示した。①は既存の Merck Index データベースを示す。②は既存の Merck Index のデータ項目を CPDO にマッピングした結果、必要に応じて細分化し、対応する CPDO に変換することを示す。③は CPDO を XML で記述して CPDO に基づく XML DB を作成し、それを持つ検索システムを ASP.NET を用いて構築したことを示す。④は、CPDO に基づく XML DB から検索者が必要とする CPDO を ASP.NET の機能を用いて取り出し、Web ブラウザにデータを送信し、表示することを示す。XML データは Web サーバ上で XSLT を利用して HTML に変換して Web ブラウザに送る。また、一部分の内容だけを表示するなどの処理が必要な場合は、その処理を VB Script を用いて記述した。

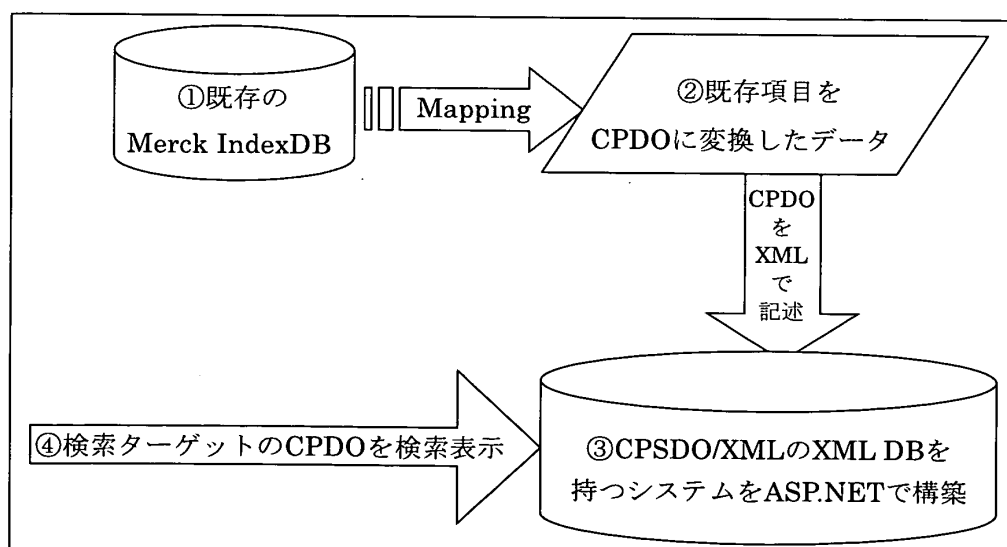


図 3: CPDO および CPSDO/XML に基づく Web 検索システムの実装のプロセス

## 4.2. 既存リソース/DBのデータ項目とCPDOの対応関係

以下、具体例を挙げて述べる。図4に、炭疽菌に効く抗生物質：CiprofloxacinのMerck Indexの冊子体の記述<sup>[20]</sup>を示す。なお、冊子体ではデータ項目名は明示されていないため、凡例に則してデータ内容を記述したものを○で囲んで示してある。Monograph numberとはMerck Indexの中の物質を識別する番号であり、その右のTitleとは物質の名称、この場合はCiprofloxacinである。これらは情報の最小単位であり、CPDOに一致する。しかし、Literature referencesはChemical InformationとBiological, pharmacological, etc. informationの2つに分かれ、さらに前者は、「Prepn(製法): 参照文献の書誌事項; 参照文献の書誌事項.」, 「Antibacterial spectrum in vitro: 参照文献の書誌事項.」等々のように主題ごとに分かれるなど内部構造を持っており、情報の最小単位とは言えない。Merck Indexは情報検索システム：DIALOG<sup>[19]</sup>にも収録されているが、そこでもLiterature referencesのデータ項目は大きな単位のままである。大きな単位のままでCPDOに対応させることもできるが、より小さい単位に分けてCPDOに対応させる方がCPDOおよびCPSDO/XMLの本来の趣旨に合っている。

図5はDIALOGに収録されているMerck Indexのデータ項目をそのままCPDOに対応させ、XMLで記述して作成したWeb検索システムの検索結果表示である。図5の中ほどの矢印で示した“LITERATURE REFERENCES”項目を見ると、全てのレファレンス文献を主題ごとには分けず、一つの項目として扱っていることが分かるであろう。Merck IndexではLiterature referencesはChemical informationとBiological, pharmacological, etc. informationの2つに分かれているが、DIALOGのThe Merck Index OnlineではLITERATURE REFERENCESのみであって分かれていない。この設計では、例えば、検索者は「炭疽菌に効く薬Ciprofloxacin」の製法だ

けを調べたい場合でも、LITERATURE REFERENCESの項目に目を通して、その中からPreparationに関連するレファレンスを抽出しなければならなくなる。これではCPDOおよびCPSDO/XMLの利点を生かしているとは言いがたい。

そこで、より小さな情報の単位、この場合は主題ごとに分けてCPDOとする設計が考えられる。図6はその設計によるシステムの結果表示である。図6の1から7までの矢印に示すように、DIALOGのMerck IndexのLITERATURE REFERENCESのデータをChemical InformationとOther Information (Biological, pharmacological, etc. informationのこと)の2つに分け、その中をさらにSubjectごとに分け、CPDOの概念に従って表示している。そのため、検索者は必要なデータ、例えば、製法(Preparation)が欲しい場合、矢印1に示す項目だけを取り出すことができる。粒度の異なるCPDOもCPSDO/XMLで記述した階層構造と対応しているため、XMLの手法：DOM<sup>[22]</sup>やXPathなどを利用して抽出することができる。

## 4.3. CPSDO/XMLのXMLスキーマ

次に、XML DBでCPDOを記述する際に使用するXMLスキーマ<sup>[21]</sup>を検討した。Merck Indexの分野に対応する既存のXMLスキーマとしては、Peter Murray-Rustによって開発されたCML (Chemical Markup Language)<sup>[23][24]</sup>および、それを科学技術及び医学分野に拡張したSTMML (Science, Technology and Medicine Markup Language)<sup>[25]</sup>がある。これらはオブジェクト指向プログラミングの観点に立ってデータ型を豊富に用意しており、また、STMMLは出版物の構造を意識した結果、大きなスキーマになっている。しかし、CPDOの視点から見ると、これらはMerck Indexの検索に必要な要素をカバーしておらず、CPDOの記述には適していない。

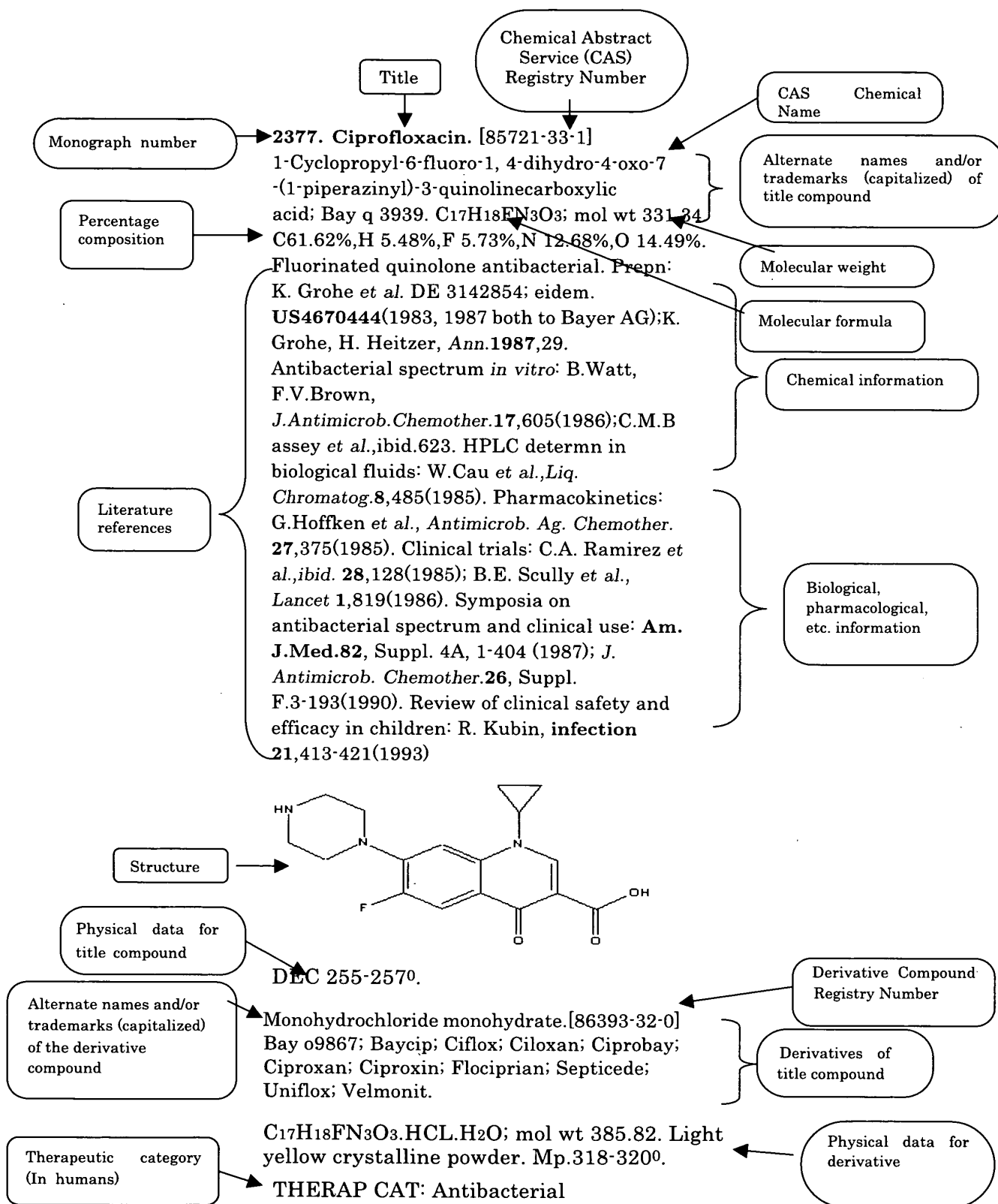


図 4: Merck Index における Ciprofloxacin の記述.

The Merck Index., 2001, p.403. Merck & Co., Inc. より引用

アドレス http://felipe4/paper/OnLineMerckindex.aspx 移動

MerckIndexデータベースにおける項目をXMLの要素に変換した結果(炭疽菌に効くCiprofloxacinを例とした)

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Merck Index                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| THE MERCK INDEX Monograph Number:2377                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Chemical Properties                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Chemical Properties, THE MERCK INDEX Monograph Number:Ciprofloxacin                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Chemical Properties, CAS REGISTRY NUMBER:85721-33-1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Chemical Properties, CHEMICAL NAME:1-Cyclopropyl-6-fluoro-1,4-dihydro-4-oxo-7-(1-piperazinyl)-3-quinolincarboxylic acid;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| SYNONYMS(Alternate names and/or trademarks (capitalized) of title compound):                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| MOLECULAR FORMULA: C17H18FN3O3                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| MOLECULAR WEIGHT: 331.34                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| MOLECULAR COMPOSITION:C61.62%,H 5.48%,F 5.73%,N 12.68%,O 14.49%                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| Sources/References                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| LITERATURE REFERENCES:Fluorinated quinolone antibacterial. Prepn: K. Grohe et al, DE 3142854;idem. US4670444(1983, 1987 both to Bayer AG); K. Grohe, H. Heitzer, Ann.1987,29. Antibacterial spectrum in vitro: B.Watt, F.V.Brown, J.Antimicrob. Chemother.17,605(1986); C.M.Bassey et al, ibid.623. HPLC determ in biological fluids: W.Gau et al, Liq. Chromatog.8,485(1985). Pharmacokinetics: G.Hoffen et al, Antimicrob. Ag. Chemother. 27,375 (1985). Clinical trials: C.A. Ramirez et al, ibid. 28,128(1985); B.E. Scully et al, Lancet 1,819(1986). Symposia on antibacterial spectrum and clinical use: Am. J. Med. 82, Suppl. 4A, 1-404 (1987); J. Antimicrob. Chemother.26, Suppl. F.3-193(1990). Review of clinical safety and efficacy in children: R. Kubin, Infection 21,413-421(1993) |
| PATENT INFORMATION:DE 3142854;US4670444                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| PATENT INFORMATION, Patent Number:DE 3142854;US4670444                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| Physical Property Information:Dec 255-257°C                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Therapeutic category (In humans):THERAP CAT: Antibacterial                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| DERIVATIVE INFORMATION                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| SUBSTANCE :Monohydrochloride monohydrate                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| DERIVATIVE CAS RN :86393-32-0                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| DERIVATIVE NAME:Bay o9867;Baycip;Ciflox;Ciloxan;Ciprobay;Ciproxan;Ciproxin;Flociprian;Septicide;Uniflox;Velmonit                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| DERIVATIVE MOL FORMULA :C17H18FN3O3.HCL.H2O;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| DERIVATIVE MOL WEIGHT :mol wt 385.82.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| DERIVATIVE PHYSICAL DATA :Light yellow crystalline powder. Mp.318-3200.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |

図 5: DIALOG における Merck Index の各項目をそのまま CPDO に対応させ、XML で記述した結果の表示

そのため我々は独自に Merck Index 用の XML スキーマ (図 7 参照) を設計した。

図 7 の 3 行目に示すように、Merck Index の XML のルート要素 (element) の名は Merck Index である。それは下位要素: Monograph の繰り返しで構成される。また、Monograph の下位の要素として、順に AccessionNumber (DIALOG の Merck Index Online ファイルのアクセス番号で、Merck Index の Monograph number に当る)、ChemicalProperties (その内容は後述)、SourcesANDReferences (内容後述) 等を設定した。ChemicalProperties の下位要素は、図 7 の 20 行目からの数行に示すように、順に MonographTitle (化学物質名で、上述の例で言えば、Ciprofloxacin)、RegistryNumber (CAS Registry Number, CAS 化学物質登録番号) 等々、MolecularComposition までが並ぶ。SourcesANDReferences の下位要素は、31 行目からの数行に示すように、ChemicalInformation, OtherInformation, PATENTINFORMATION である。最

後の要素は DIALOG で設定されているデータ項目に対応する。40 行目からの記述は ChemicalInformation の下位要素に Subject を、48 行目からの記述は OtherInformation の下位要素に Subject を、繰り返し可として設定したことを示す。55 行目からの記述は Subject の下位に繰り返し可で Ref(参考文献の書誌データ) を置けることを示す。これら個々の要素が CPDO に対応する。また、ChemicalInformation のように、要素によっては下位の要素を持たせることによって、CPDO の階層構造を実現した。

このように、CPDO に基づく XML スキーマが設計でき、これを用いることによって様々な階層に属する CPDO に対する処理が可能になる。



|                                                                                                      |                                                                     |          |
|------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|----------|
| アドレス  http://felipe4/paper/Paper/2DOCPMerckindex.aspx                                                |                                                                     | 移動   リンク |
| CAS REGISTRY NUMBER:85721-33-1                                                                       |                                                                     |          |
| CHEMICAL NAME:1-Cyclopropyl-6-fluoro-1,4-dihydro-4-oxo-7-(1-piperazinyl)-3-quinolinecarboxylic acid; |                                                                     |          |
| SYNONYMS(Alternate names and/or trademarks (capitalized) of title compound): Bay q 3939.             |                                                                     |          |
| MOLECULAR FORMULA: C17H18FN3O3                                                                       |                                                                     |          |
| MOLECULAR WEIGHT: 331.34                                                                             |                                                                     |          |
| MOLECULAR COMPOSITION:C61.62%,H 5.48%,F 5.73%,N 12.68%,O 14.49%.                                     |                                                                     |          |
| Sources AND References :                                                                             |                                                                     |          |
| Chemical Information :                                                                               |                                                                     |          |
| Generic Name:Fluorinated quinolone antibacterial.                                                    |                                                                     |          |
| 1                                                                                                    | Subject,Prepn:                                                      |          |
|                                                                                                      | REFERENCE:K. Grohe et al, DE 3142854;                               |          |
|                                                                                                      | REFERENCE:eidem. US4670444(1983, 1987 both to Bayer AG);            |          |
|                                                                                                      | REFERENCE:K. Grohe, H. Heitzer,Ann.1987,29.                         |          |
| OtherInformation :                                                                                   |                                                                     |          |
| 2                                                                                                    | Subject,Antibacterial spectrum in vitro:                            |          |
|                                                                                                      | REFERENCE:B.Watt, F.V.Brown, J.Antimicrob.Chemother.17,605(1986);   |          |
|                                                                                                      | REFERENCE:C.M.Bassey et al,ibid.623.                                |          |
| 3                                                                                                    | Subject,HPLC determ in biological fluids:                           |          |
|                                                                                                      | REFERENCE:W.Gau et al,Liq. Chromatog.8,485(1985).                   |          |
| 4                                                                                                    | Subject,Pharmacokinetics:                                           |          |
|                                                                                                      | REFERENCE:G.Hoffken et al, Antimicrob. Ag. Chemother. 27,375(1985). |          |
| 5                                                                                                    | Subject,Clinical trials:                                            |          |
|                                                                                                      | REFERENCE:C.A. Ramirez et al,ibid. 28,129(1985);                    |          |
|                                                                                                      | REFERENCE:B.E. Scully et al, Lancet 1,819(1986).                    |          |
| 6                                                                                                    | Subject,Symposia on antibacterial spectrum and clinical use:        |          |
|                                                                                                      | REFERENCE:Am. J. Med. 82, Suppl. 4A, 1-404 (1987);                  |          |
|                                                                                                      | REFERENCE:J. Antimicrob. Chemother.26, Suppl. F.3-193(1990).        |          |
| 7                                                                                                    | Subject,Review of clinical safety and efficacy in children:         |          |
|                                                                                                      | REFERENCE:R. Kubin, infection 21,413-421(1993)                      |          |
| PATENTINFORMATION :                                                                                  |                                                                     |          |
| Patent Number.DE 3142854;US4670444                                                                   |                                                                     |          |

図 6: 主題ごとに分けた CPDO に基づく Merck Index の XML データの表示

#### 4.4. CPDO および CPSDO/XML に基づく XML DB からの検索と表示

Merck Index の XML DB から検索ターゲットである CPDO を取り出した例を図 8 に示す。ここでは「炭疽菌に効く薬 Ciprofloxacin」の製法に関する特許と雑誌文献の書誌データの CPDO を、ASP.NET 上で XML スキーマと DOM を用いて取り出し、表示している。さらに、特許や雑誌の原文献リソースが Web 上に存在すれば、リンク機能を用いる連鎖検索も可能となる。

CPDO を XML データの要素として記述したため、基本的に全ての XML 関連技術が適

用できる。たとえば、必要な項目をメニュー選択により特定し、該当する CPDO を検索し、その結果を表示した例を図 9 に示す。

#### 5. 考察

既存の情報検索システム、たとえば DIA-LOG では、ピンポイントの情報を与えるとは限らず、関連の情報も含めた形で検索結果を表示するケースが多く見られる。これには次に示す 2 つの理由があると考えられる。第 1 に、既存の情報検索システムが採用しているデータベースの方式では情報に階層構造を持たせることが難しいため、たとえば、Literature

```

<?xml version="1.0" encoding="UTF-8" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified">
<xs:element name="MerckIndex">
    (中略)
</xs:element>
<xs:element name="Monograph" />
    <xs:complexType>
        <xs:sequence>
            <xs:element ref="AccessionNumber" />
            <xs:element ref="ChemicalProperties" />
            <xs:element ref="SourcesANDReferences" />
            <xs:element ref="Structure" />
            <xs:element ref="PhysicalProperty" />
            <xs:element ref="TherapeuticCategory" />
            <xs:element ref="DerivativeInformation" />
        </xs:sequence>
    </xs:complexType>
</xs:element>
<xs:element name="AccessionNumber" type="xs:short" />
<xs:element name="ChemicalProperties">
    <xs:complexType>
        <xs:sequence>
            <xs:element ref="MonographTitle" />
            <xs:element ref="RegistryNumber" />
            <xs:element ref="ChemicalName" />
            (中略)
            <xs:element ref="MolecularComposition" />
        </xs:sequence>
    </xs:complexType>
</xs:element>
<xs:element name="SourcesANDReferences">
    <xs:complexType>
        <xs:sequence>
            <xs:element ref="ChemicalInformation" />
            <xs:element ref="OtherInformation" />
            <xs:element ref="PATENTINFORMATION" />
        </xs:sequence>
    </xs:complexType>
</xs:element>
<xs:element name="ChemicalInformation">
    <xs:complexType>
        <xs:sequence>
            <xs:element ref="GenericName" />
            <xs:element ref="Subject" maxOccurs="unbounded" />
        </xs:sequence>
    </xs:complexType>
</xs:element>
<xs:element name="OtherInformation">
    <xs:complexType>
        <xs:sequence>
            <xs:element ref="Subject" maxOccurs="unbounded" />
        </xs:sequence>
    </xs:complexType>
</xs:element>
<xs:element name="Subject">
    <xs:complexType mixed="true">
        <xs:choice minOccurs="0" maxOccurs="unbounded">
            <xs:element ref="Ref" />
        </xs:choice>
    </xs:complexType>
</xs:element>
<xs:element name="Ref" type="xs:string" />
    (中略)
</xs:schema>

```

図 7: CPDO および CPSDO/XML に基づく Merck Index のための XML スキーマ

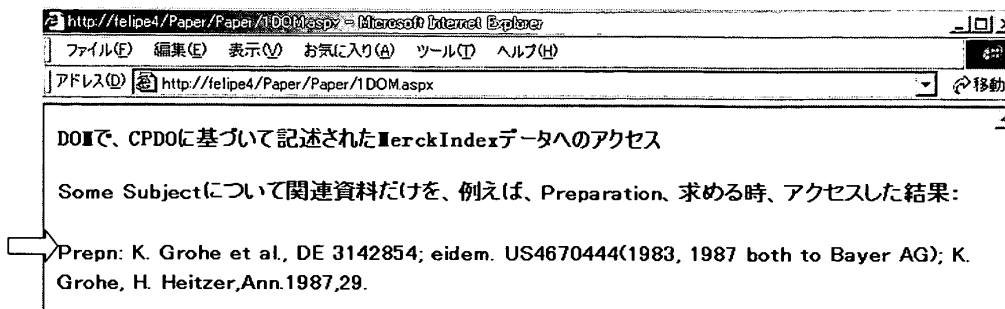


図 8: 検索対象となった CPDO(参照文献の書誌データ)のみを取り出した例

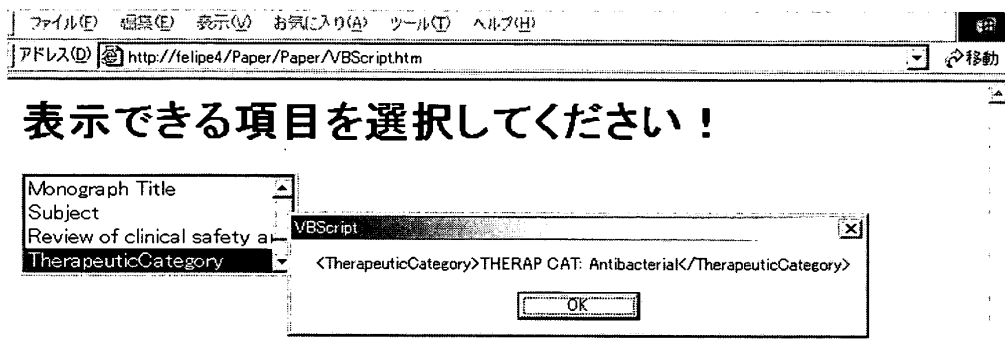


図 9: 必要な CPDO を特定した検索と表示

references, Chemical information, Subject, Reference といった階層構造を扱うよりも, Literature references の単位でまとめる方が処理が容易だからである. 第2に, 検索者は検索結果を見て判断するのであるから, ピンポイントに絞って情報を提供して検索者の意図に外れるよりも, 関連の情報を含めて提供した方が得策だからであろう.

しかし, これからの検索者は単に結果を読んで理解するだけでなく, 得た結果のデータを再構成, 再処理することもあり得るし, これにより調査・研究の効率が上がることになる. これを可能とするのが, 我々が提唱する CPDO および CPSDO であり, その XML による記述と処理システムは以上述べたように可能である.

CPSDO/XML の考え方と関連する他の手法を比較する. まずフレームとの比較では, 情報の構造化, すなわち, 階層構造とリンク関係の導入はフレームと同じであるが, CPSDO/XML は表現に XML を用いるため, 同一階層の上の要素間の順序関係も表現でき

ること, インターネット上の異なる処理系の間でデータを交換し利用できることが長所である. オブジェクトデータベースとの比較では, XML を用いることにより, 複合オブジェクトよりも容易に表現できる. また, 単に XML を用いるだけでは, オブジェクト指向の長所が活かせないが, スキーマにオブジェクト指向の考え方を取り入れた XML スキーマを用いることにより, 処理手続きを記述する際にオブジェクト指向の長所を取り入れることができた. 最後に XML と比較する. XML による構造設計の際には情報をどのように見るかによって記述形式が異なり, また, 設計の良し悪しが分かれるが, CPDO および CPSDO の考え方を取り入れることによって初めて有用な設計ができたと思われる.

CPDO および CPSDO/XML に基づくシステムの実現に問題があるとするれば, 実際に CPDO をどのようにして得るかでであろう. 考えられる方法は次の2つである. (A) 既存の情報検索用データベースあるいは情報リソースの中のデータ項目から変換プログラムによっ

て自動的に得る方法と、(B) データベースを作る段階から CPDO を意識して作るという方法である。

(A) の方法の場合、既存のデータベースあるいは情報リソースの中で、CPDO に 1 対 1 に対応するデータ項目については、プログラムによる自動変換が可能であり、問題はない。しかし、Merck Index の DIALOG 版に見られるように、現状ではデータベースとして提供されるデータ項目の粒度は必ずしも細くないため、CPDO に対応するより粒度の小さな単位に分ける必要が出てくる。Merck Index の場合は、「:」、「;」、「.」といった記号によってデータが区切られているため、かなりの程度、自動変換が可能であるが、全データが 100% 人手無しで自動変換できるとは言いがたい。

(B) の方法の場合、データベースの作成コストが上がる可能性があることが問題である。ただ、たとえば、Merck Index のデータが、「:」、「;」、「.」といった記号によって明確に区切られていることをみると、製作段階からデータの構成要素が意識され、区別されていると推測できる。また、実際に、たとえば、CAS では論文の書誌項目のうち、雑誌タイトル、巻、号、ページなどは別のデータ項目として作られている。さらに、作成コストが上昇するとしても、そのコストが何らかの方法で回収できるようになれば、その問題は解決できる。

そこで我々は、CPDO の考え方をデータベース製作段階から導入すること、情報検索などで情報を提供する際には、CPDO 単位での提供も行い、その場合は CPDO の単位での課金を行うことを提案する。これが実現すれば、CPDO の利用が進み、異なるデータベースあるいは情報リソースから集めた CPDO を組み合わせることによって、目的別のデータベースあるいは情報リソースを、従来よりも低いコストで作ることが可能となり、それがまた新しい情報活用を進めるといった好循環が現出されると期待できる。

## 6. 結論

我々は、CPSDO/XML (Computer Processable Structured Digital Object based on XML) の概念とそれに基づく検索システムを提唱した。CPSDO は、利用者が扱う情報の最小単位を構成要素とする構造化情報表現である。たとえば、もしも「検索のターゲット」すなわち「利用者が検索によって得たい情報対象」が情報の最小単位であれば、それは CPSDO の構成要素であり、もしも情報の最小単位を組み合わせただけのものであれば、それは CPSDO の一部分である。

CPSDO/XML は XML によって記述されるデジタルオブジェクトであり、XML のタグ名称は記述内容を示す。そのため、CPSDO/XML はソフトウェアによって処理可能である。また、検索対象の Web ページあるいはデータベースが CPSDO/XML に基づいて構築されれば、該当する情報の構成要素あるいはその組み合わせのみを検索でき、ノイズの少ない効率的な情報検索が可能になり、また、検索者が得たいピンポイントの情報が得られることを、典型的なファクト・データベースである Merck Index を例とするプロトタイプを構築して実証した。

本論文では、CPDO を、既存のデータベースあるいは情報リソースのデータ項目からプログラムによる変換処理によって、ほぼ自動的に得られることを示したが、この方法が全てではなく、CPDO の考え方をデータベース製作段階から導入すること、情報検索などで情報を提供する際には、CPDO 単位での提供も行い、その場合は CPDO の単位での課金も行うことも提案した。これが実現すれば、CPDO の利用が進み、目的別のデータベースあるいは情報リソースが従来よりも低いコストで製作でき、これがまた新たな情報活用を促進するといった好循環が期待できると主張した。

## 参考文献

- [1] International Standard Organization: "ISO/IEC 13250 Information Technology - SGML Applications - Topic Maps", ISO/IEC 13250, 2000.
- [2] The Semantic Web: <URL:http://www.w3.org/2001/sw/>
- [3] T. Berners-Lee et al.: "The Semantic Web", Scientific American, May 2001. (日本語訳が「自分で推論する未来型ウェブ」のタイトルで日経サイエンス, 2001年8月号に掲載されている)
- [4] Chemical Abstract Service: <URL:http://www.cas.org/>
- [5] Elsevier Science: <URL:http://www.elsevier.com/>
- [6] ISI: <URL:http://www.isinet.com/isi/index.html>
- [7] M. Minsky: "A framework for representing knowledge", in P. Winston (ed.) "The psychology of computer vision", pp.211-277, McGraw-Hill (1975).
- [8] D.G. Bobrow, T. Winograd: "An overview of KRL, a knowledge representation language", Cognitive Science, Vol.1, pp.3-46 (1977).
- [9] R.B. Roberts, I.P. Goldstein: "The FRL primer", MIT AI Lab. Memo 408 (1977).
- [10] D.B. Lenat; R.V. Guha: "Building large knowledge-based systems : representation and inference in the CYC project", Addison-Wesley Pub. Co., 1989, 360p.
- [11] STN International: <URL:http://www.cas.org/istn.html>; <URL:http://pr.jst.go.jp/db/STN/index.html>.
- [12] F. Bancilhon et al. eds. "Building on Object-Oriented Database System", Morgan Kaufmann, 1992.
- [13] D. Jordan: "C++ Object Databases Programming with the ODMG Standard", Lucent Technologies, 1998; 玉井浩 訳: 「C++オブジェクトデータベース ODMG 標準によるプログラミング」アジソン・ウエスレイ, 1998年.
- [14] Science Direct: <URL:http://www.info.sicencedirect.com/>
- [15] Extensible Markup Language (XML): <URL:http://www.w3.org/xml/>
- [16] Resource Description Framework (RDF): <URL:http://www.w3.org/RDF/>
- [17] H. Ishizuka; Y. Li: "Digital Library System Based on Concept of Computer-processable Digital Objects", Proceedings of Digital Library - IT Opportunities and Challenges in the New Millennium, p.315-326, 2002.7, Beijing.
- [18] Microsoft ASP.NET: <URL:http://www.asp.net/>
- [19] Thomson, DIALOG: The Merck Index Online: <URL:http://library.dialog.com/bluesheets/html/b10304.html>
- [20] The Merck Index., 2001, p.403. Merck & Co., Inc.
- [21] XML Schema: <URL:http://www.w3.org/XML/Schema>
- [22] Document Object Model (DOM): <URL:http://www.w3.org/DOM/>
- [23] Peter Murray-Rust, Henry S. Rzepa: "Chemical Markup, XML, and the Worldwide Web", J. Chem. Inf. Comput. Sci., 1999, Vol.39, p.928-945.
- [24] CML: <URL:http://www.xml-cml.org/information/dictionaries/index.html>
- [25] STMML: <URL:http://www.xml-cml.org/stmml/stmml.xsd>

(2002年11月19日受付)

(2002年12月25日採録)

## 「ドキュメント・マネジメント」による 知識創発型の地域社会づくり

### A Community Aims at “Knowledge-Emergent Society” by “Document Management”

西村 健\*

Ken NISHIMURA

分権型の「知識創発型社会」への移行にあたり、地域が自立的な活動を生産的に展開するための情報・知識プラットフォームとして、電子文書システムの構築をどのように進めるかが課題である。ところが、これまで自治体は、国⇒県⇒市町村とタテに繋がった官僚制原理による請負執行の体質が根強く、地域をヨコにネットワーク化し、主体的に情報・知識を生産、発信する情報政策を展開できないでいる。

地域における営みは、さまざまな複合課題を発生し、これに対処するには、地域を構成する各主体が、いかにして知識資源を共有化し、活用しつつ協働を実質化できるか、が問われる。e-デモクラシー概念に実体を与えるためにも、多角関係の協働が成立するよう情報・知識の動態を制御するXMLに基づく「ドキュメント・マネジメント」のプラットフォーム構築が重要である。

“Knowledge-emergent society” aims at that people of a community can work independently towards solution of a local subject. In order to use for activity of such intellectual production, it is a subject as a platform of information and knowledge to promote construction of an electronic document system.

However, as for a local government, a bureaucracy principle still exists persistently. A local government prefers to process the work given by several Ministries or the prefecture. Then, the local government cannot develop information policy which connects up issues of the area horizontally with a network.

In order to cope with the compound subject of a community, the member of the community, such as residents, NPO, or a company, needs to share knowledge resources, need to aim at practical use of them, and need collaboration.

Here is the important view of “e-democracy”. In order to realize the view, platform construction of XML-based “document management”, which controls the dynamics of information and knowledge, is indispensable.

キーワード：XML, ドキュメント・マネジメント, 知識創発型社会, 電子自治体,  
e-デモクラシー, 多角関係のダイナミクス

XML, Document management, knowledge-emergent society, electronic local government,  
e-democracy, multiple-related dynamics

#### 1. はじめに

情報化社会への移行にあたり、「ものづくり大国」である日本では、かねてより、ホワイトカラーの附加価値生産性の低さが指摘されてきた。この大きな課題に対処し、生産性の向上を図る上で、BPRやナレッジマネジメント

ト等の新経営手法と一体的な「XMLによるドキュメント・マネジメント」の考え方が重要であることについて、かつて本誌上で論じた<sup>[1]</sup>。

今回はこれを踏まえ、「知識創発型社会」と定義され、このところ急激に進展している本格的なネットワーク社会の到来とともに、一つの組織内だけで論じられがちであった旧来

\* 株式会社ドキュメント・エンジニアリング研究所

の生産性の考え方を拡張し、地域社会が全体として情報・知識の利活用による生産性向上を獲得していくことの重要性について論じることとする。

今日の日本社会は、いわゆる構造改革により歴史的転換を図ることが求められている。その中核をなす課題認識は、明治以後続いてきた官僚組織による中央集権体制から、分権型国家構造（地域の自立）へと本質的転換を図ることによって、真に民主的かつ自立的な活力に満ちた「知識創発型社会」が築かれなくてはならない、というものであろう。こうした分権型地域社会づくりの理念を、地域の場において実現していく上での前提として、情報・知識による価値創出という新たな生産のしくみを、どのようにして実装することが可能とされるか、という課題について考えてみたい。

## 2. 電子情報をベースとした社会づくり

「IT革命への速やかな対応」が国是とされている。

「e-Japan 戦略」<sup>[2]</sup>の中では、IT革命の意義について、次のように述べている。

コンピュータや通信技術の急速な発展とともに世界規模で進行するIT革命は、18世紀に英国で始まった産業革命に匹敵する歴史的な大転換を社会にもたらそうとしている。産業革命では、蒸気機関の発明を発端とする動力技術の進歩が世界を農業社会から工業社会に移行させ、個人、企業、国家の社会経済活動のあり方を一変させた。これに対して、インターネットを中心とするITの進歩は、情報流通の費用と時間を劇的に低下させ、密度の高い情報のやり取りを容易にすることにより、人と人との関係、人と組織との関係、人と社会との関係を一変させる。この結果、世界は知識の相互連鎖的な進化により高度な付加価値が生み出される知識創

発型社会に急速に移行していくと考えられる。

要するに、IT革命とは、情報技術の急速な発達、密度の高い情報・知識のやり取りを容易に可能とすることにより、人と組織と社会を通じた社会経済活動のあり方を革命的に転換させることだ、と説明され、転換後の社会を「知識創発型社会」と定義している。

これを実現するための重点政策の一つに「電子政府の実現」が掲げられ、その中で、人と組織と情報技術との関係について、次のような趣旨を述べている<sup>[2]</sup>。

「電子政府」「電子自治体」とは、文書の電子化とペーパーレス化による情報共有化と業務改革を通じて、2003年度までに「電子情報を紙情報と同等に扱う行政を実現」すること（e-Japan 戦略の重点政策「II. 3. 電子政府の実現」の「(2) 目標」）。

すなわち、ここで述べられていることは、

従来の行政活動では、紙文書と対面的処理をベースとすることによる業務の非効率性と情報の偏在性が際立っていた。この弊害を克服するには、紙文書と対面的処理をベースとすることから、電子情報をベースとする行政活動へと転換し、知的生産性と業務効率の向上を飛躍的に高めなくてはならない

との基本認識として理解される必要がある。これまでの紙媒体に依存した活動のあり方から、電子媒体へと移行することによって、始めて革命的な転換が可能となるからである。

ところで、「同等に」の意義に関して、しばしば「電子情報と紙情報とを一緒に扱う」との誤解がなされるが、正しい読み方を記述するなら、

行政活動がこれまで紙情報を扱って行われていたのと同じように、これからの行政活動が、電子情報を扱って行われるようにする、ことを実現する

という意味に理解される必要がある。そして、こうした誤解が発生しやすい土壌から、次に掲げるように、自治体情報政策が画一的な陥穽にはまりやすい状況が産み出されている、と考えられる。

### 3. 電子申請に偏る情報化政策

さて、IT革命の本質的意義理解に添った重点施策として電子政府が打出され、電子文書をベースとする行政活動への転換が掲げられたところまでは、当を得たものと理解されるが、IT革命の実体を構築すべき地域の場において、現実に進められている情報政策はいかなるものであろうか。

結論から言えば、残念ながら、IT革命の本質的意義理解に添った方向には程遠い、と評すべき実情にある。

それは、第一に、せっかく電子政府が「電子文書をベースとする行政活動への転換」という正しい方向性を掲げておきながら、現実には、電子申請が前面に押し出され、矮小化されていることである。むしろ、電子申請がもたらす国民にとっての利便性向上と業務処理の効率化に関する貢献を否定しようというものではない。にもかかわらず、電子申請を大目標に掲げるだけでは、旧来の手続処理（ルーチン・ワーク）の効率化を求めるだけで、新たなIT革命の本義である知的創発性に貢献するところまでには至らないのである。

そして、第二に、より本質的な事柄として、主体的受けザラとなるべき自治体の情報政策＝電子自治体が、こうした国の電子政府＝電子申請を中核とする情報政策の受け売りに終始している状況が根深く認められる、ということがある。

こうした状況にあって、なぜそうなっているのか、また、推進されるべき情報政策の方向性とはどうあるべきか、ということ考察し、論を進めるについて、少し回り道ではあるが、背景を探り、事態を明確化しておくこととする。

### 4. 根強い執行型行政観

そもそも分権型国家構造を前提とする高度情報社会＝知識創発型社会とは、地域が自主的、自立的にネットワーク連携しつつ、コラボレーションが実体化できることを目指すものであるべきところ、その基盤を構築すべき情報化政策が、相変わらず国の省庁のタテ割分担による政策の延長、執行として位置付けられるところに本質的矛盾を抱えている。

このことについて、次の図1により考えてみよう。

国とは、社会経済活動の諸相に合わせたセグメントをそれぞれ所管する各省庁が、タテワリに存在しており、これまでの自治体は、各省庁から別個に政策と業務、予算を受け取り執行してきた。各省庁が政策と業務、予算を別個に下ろしてくる以上、受け取る自治体側としても、各省庁に対応するタテワリ組織が別個に処理することが通例で、自治体内部といえども、ヨコに独自に括り直し連携性をとるということは、ほとんど行われてこなかった。

例えば、図中のA省が県庁のa部と矢印で結びつき、さらに市町村のa課と結びついて、そこから始めて住民との関係性が具体化する、ということがこれまでの業務処理の流れであった。電子申請とは、こうしたタテの

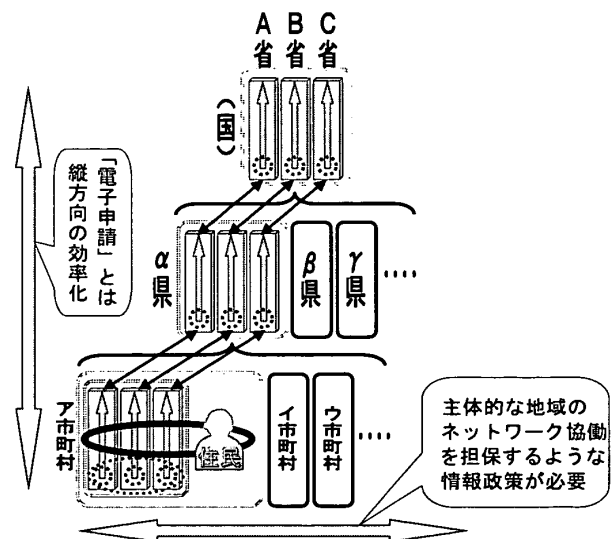


図1: 国県市町村の縦の繋がり地域社会の横の協働ネットワーク



業務処理の流れを効率化しようとするものであるといえよう。なぜ知識創発へ至る情報化を掲げているながら、電子申請だけが一人歩きしているかのような現状が広まっているのであろうか。

これまでの近代社会における発展の姿を省みると、主として画期的な動力やエネルギー革命とともに描かれてきたと思われる。機械化、システム化といえ、力仕事を大幅にスピードアップし省力化するというイメージである。同様に、コンピュータ・システム、情報技術に求められてきた役割とは、事務処理における人力による煩雑な手間仕事（力仕事）を大幅に削減することで、行政事務に当てはめれば、まさに煩雑な手間にかかる申請事務や手続処理の業務が大幅に簡便化、迅速化される、という話が一般に理解されやすい、ということであろう。この図式化により明らかとなるのは、行政事務を始めとする事務処理とは、ルーチンとして定められた煩雑な手続処理にかかる手間作業のこと、という思い込みの存在である。さらにいえば、これらの業務処理やその根拠たる政策、施策とは、中央の官僚や組織の上層部が決定し、与えられたもので、一たん与えられたからには、いかに効率的に処理するか、に集中すべきだ、といった官僚主義組織原理から来る根強い執行型の行政活動観（コンプライアンス原理）に由来する、とも言えよう。

## 5. e-デモクラシーの登場

これに対し、分権型国家構造を目指す地域の情報政策としては、地域が自立し、真に活力ある地域社会へ向けて発展していけるために、ITが貢献できる方策を講じるものでなくてはならない。あくまで地域における主体的な政策の形成とまちづくりへ向けた積極的活動を支援することに主眼が置かれなくてはならない。これが「e-デモクラシー」として最近ことに強く説かれている事柄の根本にある考え方であろう。e-デモクラシーとは、地方

自治がこれまで行政主導で語られてきたことを転換し、住民が地域コミュニティづくりに主体的に参加していけるよう、積極的にその活動を支えるインフラとしてIT、ことにインターネットを活用することを意味している。

その前提には、地域コミュニティの活性化を目指し、公民パートナーシップを確立する、という考え方の抜本的転換と行政構造の改革が伴わなくてはならない。

すなわち、平石正美<sup>[3]</sup>によれば、考え方の根本は、

- 地域のさまざまな課題の解決を、コミュニティ、ボランティア、NPO、企業など行政外部のアクターに求める  
それには、
- 協働を高め、アカウントビリティを向上させ、課題解決へ向け外的な資源活用の可能性を高めることによって、住民の参加形態やガバナンスの仕組を有意義なものに変革する

とされている。

では、以上のことから結論して、地域の情報政策とは、端的に「電子申請」から「e-デモクラシー」への進化が図られれば、あるべき電子自治体の実現される、といえるであろうか。

むしろe-デモクラシーが目指されなくてはならない、という本質的意義に間違いはないのであるが、現状の議論を見ている限りでは、電子申請すなわちたんなる事務処理の効率化、を目指すスタンスに対するカウンター主張としてe-デモクラシーが掲げられているようにも見受けられる。肝心なことは「住民と行政とのコラボレーションを支えるための情報・知識インフラ整備」という課題に、いかに実体を与え、実質的に取り組んでいくか、である。ところが、こうした大問題には、ほとんど焦点が当てられていない。実際に行われているところを観察すると、住民電子会議室の設置とアドバイザーによる議論の集約及び若干の情報提供といった程度をもって、e-デモクラシーと称しているケースが認められる。

## 6. 協働を支える情報・知識インフラづくり

「住民と行政とのコラボレーション」を進めるにあたり、一括りで「住民」としてしまうのではなく、行政との関係を少なくとも次の三つの態様に分けて、それぞれの特性を考慮することが求められる。

- ① 地域社会を構成する主権者としての住民との関係
- ② お客様（サービスの受け手）としての住民との関係
- ③ ビジネス・パートナーとしての住民・地域組織（NPO、企業等）との関係

①についてみれば、「地域社会を構成する主権者としての住民との関係」をテーマとする分野が、まさに「e-デモクラシー」という発想に結び付く。

②についてみれば、「お客様（サービスの受け手）としての住民との関係」をテーマとする分野とは、CRM（ワンツーワン・マーケティング）として最近のマネジメント論で盛んに取上げられている事柄に結び付く。

③についてみれば、「ビジネス・パートナーとしての住民・地域組織（NPO、企業等）との関係」をテーマとする分野とは、これも新たな経営モデルとして取上げられるアウトソーシング論の考え方（すべてを内製化するのではなく、得意分野を連携活用し全体最適化を図る）に結び付くものである。

これらは、それぞれに情報化社会に対応した新たなマネジメント原理に立脚したアプローチをとるべきことを示している。そうした手法が様々に提起されており、その成果を巧みに取り入れつつ推進していくことが求められる。そのためには、住民側も、行政側も主体的な存在へと変わらねばならない。ここで特に問題となるのが、自治体側の体勢である。

地域経営のセンター機能を支える自治体組織が中核機能を発揮すべきであるが、その機能発揮のあり方自体が、従来型の閉じた系から、地域社会に開かれた系へと転換しなくて

はならない。ここに難題が存する。

先に図-1に関して説明したとおり、これまでの自治体の体質は、省庁からの請負執行型のタテワリ官僚組織原理がしみわたり、庁内でヨコに連携する発想は脆弱である。それとともに、価値を産み出す仕組みが、地域という場及び住民や個客といったサービスの実需者にシフトした、という従来とは逆転的な情報化社会の本質にかんする如実な認識が、まだ希薄なのである。市町村、ことに現場に近い存在である小規模自治体にあっては、地域という場からの発想に近づくことは比較的容易であろう。しかるに県庁にあっては、図-1で描いたように、元来の位置付けが国（各省庁）からの施策、事業を市町村へ「下ろす」という中間媒介者の機能体として存在し続けてきた重い経緯がある。それゆえ、三重県のようにトップからの強い方向付けがなされれば別であろうが、存立の根本を、住民のため、あるいは市町村のため、と逆転するところからすべてを再構築し直すことと一体的に情報政策を考える、という重い課題への道は、限りなく困難というべきであろう<sup>[4]</sup>。

さて、この難題に対処するにあたり、世に行われている一般の論調では、組織風土及び職員の意識改革の側面ばかりを取上げている。これでは全くIT革命には至らない。ここに新たな「和魂洋才」の考え方によることが重要とされる。

かつて日本の変革期において採用された社会システム改革の方法を「和魂洋才」と名付けている。「和魂」とは、困難課題の解決へ立ち向かうには、他人頼みではなく、自分自身で独創的な考えを工夫創出して対処する気概がなくてはならない、ということである。「洋才」とは、「和魂」を支え、その創意を現実化するについては、観念的な竹やり精神主義に陥ってはダメで、進取の気性に基つき、合理的な科学技術の成果を柔軟に採り入れ、新たな生産的システムを構築することにより一体的に推進すべし、というシステム思考を指していると考えられる。

## 7. 新マネジメント論と一体的な情報政策

今日における「和魂洋才」とは、価値を産み出す元が、モノの産出から知識の産出へとシフトしていくため、第一に、組織における活動の姿が、これまでの官僚組織原理によるものから、知識を産出することを旨とする「新たな組織運営」へと転換する必要がある。

それとともに、第二に、知識を媒介する文書内容を縦横に駆使するため、電子媒体の柔軟な処理のための「新たな情報システム・技術」の導入により担保する必要がある、ということになる。

このような「新たな組織経営論」と「新たな情報システム・技術」との一体的、革新的な導入がなされてはじめて、自治体が新たな主体的役割を發揮することが可能となる。ところが、行政の現場には官僚制原理と一体的な文系と理工系の分断現象が根強く存在している。このことが、これまでのコンプライアンス中心から、次のようなパフォーマンス主導への転換にあたり、自治体改革の躓きとなっていると考えられる。

「不確実性」の時代にあって、状況の変化に柔軟に対応するため仮説設定と現実的な解決方策が求められ、自治体マネジメントは、「計画」から「戦略」へシフトすることが不可避となる。それは次に掲げるような組織活動の新たな性質を特徴とすることから、担当者は、これまでのたんなる執行者から、それぞれ次に示すような情報編集活動を主体的・能動的に演じていくプロデューサーないしはコーディネータとしての性格を發揮する者へ進化していくことが求められる。

- 戦略的対応のための仮説設定（目標設定）  
⇒ 現場情報に基づくマイクロ・マクロループ
- 新たな政策課題は境界領域に発生  
⇒ セクションを跨る専門知識を糾合、組合せ編集
- 行政内部と地域社会とを繋ぐ  
⇒ 協働の前提たる情報・知識コンテキスト共有

このような多角関係のダイナミクスを特性とする今後の地域づくり活動を支えるためには、情報・知識を媒介する文書の扱いは、旧来の定型的処理から、発想、推論、評価など意味概念を扱うことを主とする方向へ変化する（図2参照）。

ところでいずれの自治体においても、電子文書のシステム化に取り組んでいるが、以上のようなプロアクティブな活動性を担保するため、情報・知識をダイナミックに利活用することには、全く焦点が当てられていない。庁内で組織を横断して情報・知識を活用し、新規課題への対処や政策の立案、評価が総合的、横断的に図りうるようにするとともに、住民と行政担当者とが主体的に協働できるためには、必要とされる様々な現場において、情報・知識の編集利活用ができることを担保しなくてはならない。

それには情報・知識がヨコに連携できること、及び常に内容が更新され、コンカレントな動態として扱えることが要求される。

例えば、新たなマネジメント論による自治体活動の主要な柱である情報公開についてみ

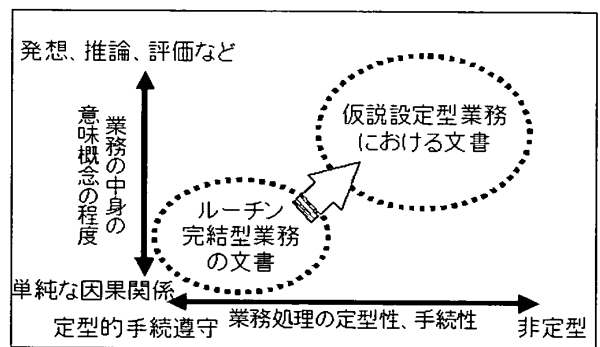
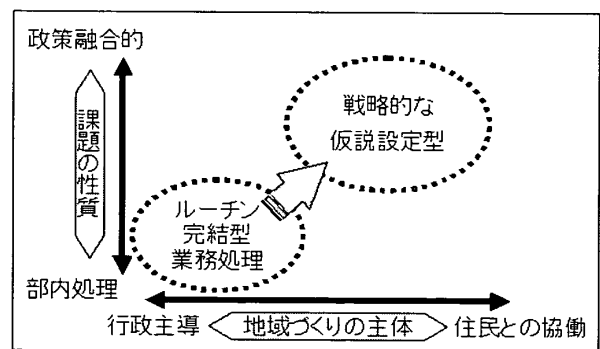


図2: 自治体業務の変化と文書の性質

ると、先にも指摘したように、これまではコンプライアンス原理を表す公文書開示（執行結果文書の公開）が中心であった。この根深い行政執行観により、これまで自治体で取組まれた電子文書システム化とは、従来から受け継がれた紙文書の記録管理の仕組みをそのままにして、たんなる公文書の電子化に止まっている。

これまでのコンプライアンス原理主導から、行政改革の目的たる成果重視原理を体現する「パフォーマンス・アカウンタビリティ」へのシフトを実体化するには、以上に考察した行政活動のダイナミクスを記述する仕組みの実装が不可欠なのである。

それにはXMLに基づくドキュメント・マネジメントの標準プラットフォーム構築が重要である（図3参照）。

以上をまとめると、自治体の新たなマネジメント原理に立脚した主体的な情報政策の中核をなすべきは、地域での創発的な協働が成り立つ前提として、情報・知識が多様な場において繋げられ、編集利活用が可能な「ドキュメント・マネジメント」という標準プラットフォームの構築である（図4参照）。

すなわち、地域に存在するNPO、企業等の組織体や住民各人、あるいは近隣自治体とのコラボレーションが成立するように情報・知識の統合的利活用が図られることが目指されなくてはならない。

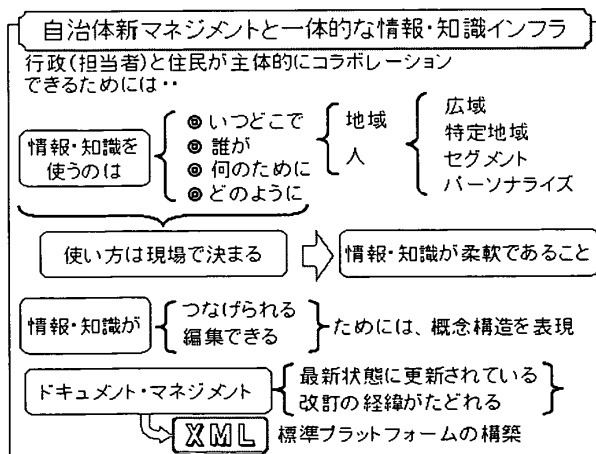


図3: 協働に不可欠な情報・知識のダイナミクス制御

それには、自治体組織内の情報・知識をネットワーク化する視点と、地域社会をネットワーク化する視点を統合化するアプローチが重要である。

これらは、情報化の対象分野として、「行政情報化」と「地域情報化」という区分けに相当しており、これまで情報化政策がそれぞれ旧総務庁と旧郵政省とに分担され、別個にシステム化されてきた。しかるに、今日求められているのは、旧来の「中央⇒地方」という図式による官僚主導の上命下達式リニア構造から、自立的な地域間のネットワーク連携方式への抜本的転換を図ることによって、分権型の「知識創発型社会」を実現することである。

合併や広域市町村圏の業務協同化という再編成も進んでいる中であって、電子自治体を推進するプロジェクトがメインターゲットとして取り組むべきは、個別の既存業務ごとのシステム化ではなく、地域における情報・知識のネットワーク活用が縦横に可能とされる仕組みの構築である。

また一方では、ウェブを活用した従来にはないサービス分野への取組みも進められている。先に、住民電子会議室の設置とアドバイザーによる議論の集約及び若干の情報提供といった程度をもって、e-デモクラシーと称しているケースが多く認められる、と指摘した

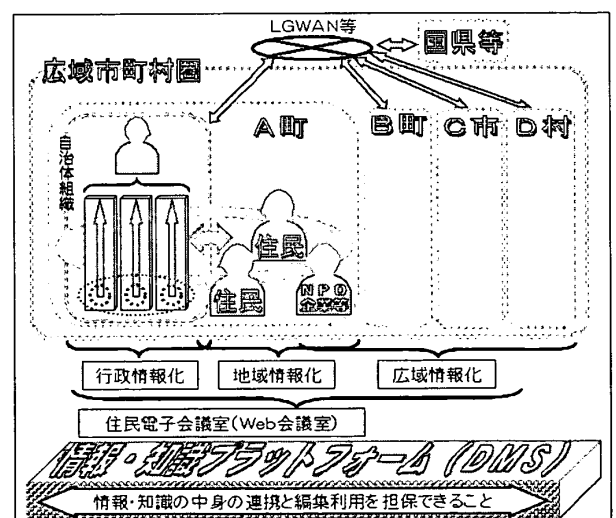


図4: 情報・知識プラットフォーム

が、ここにもはまりやすい陥穽が口を開けている。

情報・知識の内容を容易に交換，連携，編集できるための方策として，XML 連携データベース機能による「ドキュメント・マネジメント」のプラットフォームを早急に標準整備することが肝要なのである。

## 参考文献

- [1] 西村健：「自治体マネジメント改革とXML」，情報知識学会誌，Vol.10，no.3，pp.27-32，2000.
- [2] 高度情報通信ネットワーク社会推進戦略本部：「e-Japan 戦略」，2001年1月22日，<http://www.kantei.go.jp/jp/singi/it2/kettei/010122honbun.html>（参照 2002-12-15）.
- [3] 平石正美：「電子政府と電子デモクラシーの諸相」，行政&ADP，Vol.38，no.4，2002.4.
- [4] 中野雅至：「電子自治体と政策理念」，行政&ADP，Vol.37，no.10，2001.10.

(2002年11月5日受付)  
(2002年12月19日採録)

## 伝統的知識と知的所有権 — 解説

Traditional Knowledge and Intellectual Property Rights:  
A Review

名和 小太郎\*

Kotaro NAWA

伝統的知識 (TK) は先住民の社会, 地域社会がもつ知識, 信条などの総体であり, 生物多様性の保全およびその持続的な応用と密接に結びついている. TK は世代から世代へと伝承され, 固有の発明者をもたない. いっぽう, 知的所有権 (IPRs) は個人によって創作された新しい知識を保護するものであり, 集合的な権利にはなじまない. したがって, TK に対する IPRs の関係は多くの国際的なフォーラムで論争を生じている.

Traditional knowledge (TK) is defined as a cumulative body of knowledge and beliefs of indigenous and local communities and closely linked to the conservation and sustainable use of biodiversity. TK is handed down from generation to generation, and has no identifiable inventor. Intellectual property rights (IPRs), in contrast, protect new knowledge that is created by individuals and do not recognize collective rights. Therefore, many controversies over the impacts of IPRs on the TK have been raised in various international forums.

キーワード: 知的所有権, 伝統的知識, 生物多様性条約, 世界知的所有権機関, ユネスコ  
Intellectual property rights; traditional knowledge; Convention on Biological Diversity;  
World Intellectual Property Organization; UNESCO

## 1. 問題の所在

## 1.1. 生物多様性条約

生物多様性条約 (Convention on Biological Diversity = CBD) は 1992 年に採択された. その目的はつぎのように示されている.

- (1) 生物の多様性の保全.
- (2) その構成要素の持続可能な利用.
- (3) 遺伝資源 (genetic resources = GR) の利用から生じる利益の公正かつ衡平な配分.

世界知的所有権機関 (World Intellectual Property Organization = WIPO) は, GR の利用について二つの可能性を示している. その一つは本来的な利用, もう一つは非本来的な利用である [WIPO, 2001]. この本来的な使用

に関係する知識が伝統的な知識 (traditional knowledge = TK) と呼ばれるものである. TK について, CBD は「生物の多様性の保全および持続可能な利用に関連する伝統的な生活様式を有する住民の社会および地域社会の知識, 工夫および慣行」という表現で言及している. つまり, TK は GR に付随した知識として示されている.

## 1.2. TK のヴァリュ・チェーン

WIPO は TK の非本来的な利用について, つぎのようなヴァリュ・チェーンを示している.

- 「(1) GR およびこれに付随する TK」 →  
 「(2) 遺伝学およびバイオインフォマティクス」 →  
 「(3) バイオ技術の研究開発」 →  
 「(4) 商業的な製品/サービス」  
 ここで (3) と (4) にかかわる知識が知的財

\* 国際大学グローバル コミュニケーション センター  
Center for Global Communications, International  
University of Japan

産権 (intellectual property rights = IPR) という型になる。なお、このモデルにおいて(1)のなかでのみ維持される知識がTKということになる。

CBDにもどる。この条約は豊かなGRを保有するが低い技術集積しかもたない国(以下、GR保有国)と、貧しいGRしか保有しないが高い技術集積をもつ国(以下、技術集積国)があることを前提にし、前者から後者への資源提供と、後者から前者への技術移転——利益還元を含む——について、両者の均衡を求めたものである。これは、上記のヴァリュチェーンにおいて、(1)のTKという知識システムに、(3)および(4)のIPあるいはIPRという知識システムを組み合わせることを意味する。ただし、TKという概念はあいまいである。このために上記二つの知識システムを組み合わせることは難しい。したがって、ここでは、まずTKについてその概念を整理し、ついでTKとIPRという二つの知識システムについてそれらの特性を比較してみたい。

## 2. TKの定義

### 2.1. CBDへのWIPOの注

TKは多義的な存在である。まず上記CBDの記述について、世界知的所有権機関(World Intellectual Property Organization)はつぎのような注釈を付けている [WIPO,2002a]。

- (1) TKは、自然と密な接触のなかで生存する世代を通して、人の集団によって作られた知識の総体である。
- (2) TKは、分類のシステム、地域環境に関する経験的な観測のセット、資源利用を制御する自己管理のシステムを含む。
- (3) 工夫 (innovation) は先住民および地域共同体の特徴であり、これによって伝統がフィルタとなり、つぎの工夫を作り出す。これは研究と応用に関する伝統的な方法であるが、そのまま存続する知識の固有な部分になるわけではない。

- (4) 慣行 (practice) は知識と工夫の表出と見るべきである。

なお、WIPOはTKを「遺産のサブセットであり、そのなかに先住民の知識 (indigenous knowledge) とフォークロアの表現を含む」と説明している [WIPO,2001]。この定義をIP制度の側から見ると、特許権に対する「知識システム」と著作権に対する「文化的な表現」とに分かれる。この報告においては、前者を扱うことにしたい。

### 2.2. UNESCOの定義

国連教育科学文化機関 (United Nations Educational, Scientific and Cultural Organization = UNESCO) は1999年に「TKと先住民の文化の表現 (expressions of indigenous cultures = EIC) に関するシンポジウム」を開催した。この報告はTKとEICについてつぎのような関連語を示している [WIPO,2002b]。

精神性、精神的知識、倫理、道徳的価値／社会制度／舞踏、儀式、儀礼的な演技および慣行／ゲーム、競技／音楽／言語／口承の物語における名前、構想、仕来り、音楽／土地、海、空／文化的意味をもつすべての場所、文化的不動産、これらに結びつく知識／文化的・環境的資源／伝統的な保全方法を含む伝統的な資源管理／すべての物質的な対象と文化的な動産／本来の地域の外で蒐集として保管されているすべてのTKとEIC／先住民の先祖の遺跡、人間の遺伝的物質／科学的・農業的・技術的・生態的知識、この知識を使うために必要な技能。

同年、UNESCOは国際学術連合 (International Council of Scientific Unions = ICSU) と共同で科学に関する世界会議を開催し、ここで『科学計画—行動の枠組み』という文書を発表している [UNESCO,1999]。この文書はTKを「現代科学とは異なる知識システム」と定義し、これに「自然界に関する経験

的な知識は現代科学に独占されるものではない。伝統的な社会は、その多くは先住民によるものであるが、かれら自身の知識のシステムを育成し、洗練してきた。これらは世界では莫大な富であり、異なった生活方式に関するパラダイムである。それは人類の遺産の一つとして認識され保全されなければならない」という説明を加えている [UNESCO,1999]。ただし、ICSUはこの定義に対して、これでは占星術や創造主義まで含んでしまうのではないかという意見を述べている [Dickson,1999]。

## 2.3. その他の定義

### 2.3.1. TK 保全側の定義

TKの定義は、その保全側と利用側とで異なる。前者については、世界保健機関 (World Health Organization = WHO) の「伝統的医学の研究と評価に関する方法論に対する一般的なガイドライン」(2000)のつぎのような定義がある [WIPO,2002b]。「知識、技能、慣行の総体であり、それらは多様な文化に固有の理論、信仰、経験に基づく。それは記述可能であるか否かを問わない。また、健康の維持、つまり、肉体的、精神的な病気に対する予防的診断、改善、処置に用いられる」。ここには「補助的／代替的／非通常的医療という用語は、伝統的医学と同義語として使われる」という注が付けられている。

また、国連食糧農業機関 (Food and Agricultural Organization = FAO) は、漁業研究諮問委員会の資料において、TKにつきのような説明を与えている [WIPO,2002b]。「世代から世代へと、とくに口承で伝えられる文化の要素をもつ」「世代から世代へ、人々によって途切れることなく引き継がれる思想または行動の様式。つまり習慣あるいは慣行」「現在に影響を及ぼす先例の固まりとして理解される一組の習慣と慣行」。

### 2.3.2. TK 利用側の定義

世界貿易機関 (World Trade Organization

= WTO) は「地域的、先住民の共同体の TK に対する IPR の保護に関する提案」(1999)において TK につきの定義を示している [WIPO,2002b]。「工夫、創造および文化的な表現であり、その現在の所有者によって生みだされ、維持されているものである。現在の所有者は、個人あるいは全共同体、自然人あるいは法人として定義され、確認されており、かれらは権利の保有者である。その所有者に対する TK の経済的、商業的、文化的な価値は、この知識が IP の主題として認識されるという法的な利益を保証し、正当化するものである」

世界銀行 (World Bank) は TK についてつぎのような説明を与えている [WIPO,2002b]。「地域的な知識である」「個々の文化または社会に固有のものである」「農業、保健、食糧備蓄、教育、自然資源の管理における地域的な意思決定について基礎となる」「個人より共同体によって維持される」「暗黙の知識であり、コード化が困難。共同体の慣行、制度、関係、儀式に埋め込まれている」

### 2.3.3. TK の特性

以上の定義をまとめると、TK はおよそつぎのような特性をもつものとして要約できる [高倉, 2001; 寺嶋, 2002; Khalil,1995; WIPO, 2001]。ただし、これらの特性のなかにはあい矛盾するものもある。

- (1) 世代から世代へと伝達される。
- (2) 特定の人びと、特定の地域に共有される。
- (3) テキストではなく、口承で伝わる。
- (4) 試行錯誤的な経験の集積である。
- (5) 時間的、循環的に変化する。
- (6) 特定集団内に秘匿される場合もある。
- (7) 信念、信仰、世界観までも含む。
- (8) 宗教的かつ代替医療的な機能をもつ。



### 3. TK の知的所有権化

#### 3.1. 事例

WIPO の示したヴァリュ・チェーンにおける TK の非本来的利用について、その事例を紹介しよう。

〔事例 1〕ターメリックはインドにおいて「お婆さんの治療薬」として傷口の処置に使われてきた。ミシシッピ大学医学センター (University of Mississippi Medical Center) は 1995 年に「傷口治療へのターメリック使用」という米国特許 (USP 5,401,504) を取得した。インド科学産業研究協会 (Council of Scientific and Industrial Research of India) は 1996 年に米国特許商標庁 (US Patent and Trademark Office = USPTO) にこの特許の無効を申し立て、それに成功した [Dutfield,2000]。

〔事例 2〕アヤフアスカはエクアドルにおいてシャーマンにより幻覚を起こす「魂の酒」として宗教的な儀式に使われてきた。米国市民の L.S.Miller は 1986 年にこの品種について米国植物特許 (USPP 5,751) を取得した。国際環境法センター (Center for International Environmental Law) は 1999 年に USPTO にこの特許の無効を申し立て、それに成功した [Downes,2000]。

第三世界の研究者には、技術集積国の企業によるこのような行為を「バイオパイラシ」として非難するものがある [Shiva,1997]。

#### 3.2. GR 保有国の主張

アンデス共同体 (Andean Community) は上記のような技術集積国の企業行動に対して 1996 年に「遺伝資源へのアクセスに対するコモン・システム」を採択している。アンデス共同体はボリビアなど南米 5 国が設けた地域協定である。

アンデス共同体はその「生物学的、遺伝的遺産および TK に関する決議 486」においてつぎのように主張している [WIPO,2002b]。「加盟国はつぎを保証しなければならない。IP の

要素に対する保護は、生物学的、遺伝的な遺産、それは先住民の TK に結びついたものであるが、それへの保護と尊重に一致するものでなければならない。したがって、その遺産、あるいはその知識から得られた材料を基礎にして開発された発明に関する特許を認可するためには、国際的な法、アンデス共同体の法、および各国の法に調和して材料を獲得しなければならない。加盟国はその集合的な知識について、先住民の、アフリカン・アメリカンの、地域的な共同体の権利と権威を認識しなければならない」。このような運動はたとえばアフリカ統一機関 (Organization of African Unity) も進めている。

#### 3.3. TK の特許適格性

##### 3.3.1. 非適格性

TK は現行の特許法になじみにくい特性をもっている [山名, 2002]。したがって、そのままでは特許権を得ることはできない。この理由はつぎのようなものである。

- (1) 特許法によれば発明者は「自然人」でなければならない。しかし、TK は集団のつくり出したものである。
- (2) 特許法によれば発明は「産業上の利用可能性」を持たなければならない。いっぽう、TK には医療に関するものが多いが、医療行為は産業には含まれない<sup>(1)</sup>。
- (3) 特許法によれば発明は「公然となったもの」ものであってはならない<sup>(1)</sup>。しかし、TK はテキストとして記述されることはなく、くわえてクラブ財として公然と非公然の中間にある。
- (4) 特許法によれば発明は「非自明性」をもたなければならない。しかし、TK はつねに改良および保守行為のなかにある。

これらの理由があるために、TK の保有集団がその TK をそのままの形で特許として権利化することは難しい。

<sup>(1)</sup>米国法は医療行為に関する制限を弱く、また、公然として認める条件については自国中心に規定している。

くわえて、特許権は新しい技術を応用前に発明に与えるものであり、その技術がもたらす社会的な影響までを評価の対象にするものではない[名和, 2002a]。さらに、TKは人類学にとっては発見の対象となる「事実それ自体」であり、IP制度はそのような事実を権利化することができない[名和, 2002b]。

### 3.3.2. 非適格性の迂回

技術集積国の企業は、そのままでは特許に非適格性をもつTKを加工し、それに特許権を設定することができる。また、これを実行している。具体例は3.1に示したとおりである。かれらは上記の不特許理由をつぎのようにして回避している。

- (1)' 現行の特許制度においては、「人為的」なものは特許の対象になりうる。たとえば天然物に対しては特許はとれないが、天然物から抽出したものは特許の対象になりうる。このような人為的な操作をおこなった者が発明者になりうる。
- (2)' 医療行為に役立つというTKは、ここに医薬的な効果をもつ物質が含まれているという事実を示している。医薬であれば特許の対象になりうる。
- (3)' TKがテキストとして記述されていないということは、TKの応用にすぎない発明に対して先行技術なしという評価を導く。したがって、その発明は特許の対象になりうる。TKがクラブ財であれば、この評価はさらに確実なものとなる。
- (4)' TKそれ自体には非自明性はないが、ここから導かれた発明に対しては非自明性を認めることもありうる。

## 4. 国際的枠組みの再設計

### 4.1. 基本的な考え方

IPに関する国際的な流通については3つの考え方がある。

- (1) IPを自由な貿易財とする。

- (2) IPを貿易財としては無効にする。この無効化は「人類の共同遺産」つまり「公有」という理念で実現する。
- (3) IPの貿易で得られた利益を関係者で分配する。

それぞれの理念は、(1)についてはWTOの「知的所有権の貿易的側面に関する協定」(Agreement on Trade-Related Aspects of Intellectual Property Rights = TRIPS, 1994)で、(2)についてはFAOの「植物遺伝資源に関する国際的申合せ」(International Undertaking on Plant Genetic Resources = IUPGR, 1983)で、(3)についてはCBDで実現している。現在は、(1)のTRIPSが圧倒的な力を持ち、(3)の具体化については第三世界の諸国による若干の試みがある[Dutfield,2000]。

詳しくいえば、IUPGRとCBDはIPのものではなく、IPをそこから引き出す「遺伝素材/遺伝資源」、あるいはTKに関する規定となっている<sup>(2)</sup>。

つまり、国際的な取り決めにおいては、素材、資源という有体物である場合、あるいは、「IPに関する権利」という形式化された情報である場合には管理対象になるが、その中間にあるTKという非形式的な情報である場合には管理対象から外されてしまう。これはTKが技術集積諸国の主導によるIP制度では処理できないという事実を示す。

参考のために示すと、特許の国際出願を国別で見よう。その件数は1997年において、米国、西ヨーロッパ諸国、日本のみで93.9%にのぼる。つまり、現行の特許制度は技術集積諸国に占有されていることになる。

### 4.2. TKのIPR化

TKを、現在のままIPR化しようとした場合、つぎの点について、国際的な合意が必要

<sup>(2)</sup>TRIPSは、(1)診断・治療・外科的方法、(2)動植物(除く、微生物)、(3)動植物生産のための本質的に生物学的な方法(除く、微生物学的方法)、に対する発明を特許の対象から外している。

となる [WIPO,2002a].

- (1) TK の IP 保護の目的と原則.
- (2) 公的な IPR 制度と慣習的な地域的な法的システムとのインタフェース.
- (3) TK に有効な創造, 所有権, 管理資格の共有.
- (4) 国境を越える地域的な TK に対する法的, 管理的な課題.

これらの課題の解決のために, たとえば TK に関するデータベースを設けようという活動もある [Kelowna,2002]. これは上記 (3)(4) 項にかかわる提案である. 問題は, このデータベースを公開すれば, それは技術集積国の企業に利益を与えるのみであり, したがって, アクセスのためのパスワードをだれに付与すべきかという問題も, TK 保有者の側からは申し立てられている.

また, 企業の GR などに対するアクセスの可否を GR 保有国の政府に判断させようという試みもある [Schiermeier,2002]. これは上記 (1)(2)(3) 項に対する提案である. 問題は, GR 保有国の政府が, このような試みを実現できる法的なインフラストラクチャをもっているかということにある.

これらの点をめぐって, GR 保有国と技術集積国とのあいだには, 深刻な政治的対立が生じている [大塚, 2002].

### 4.3. まとめ

この報告で紹介した多様なドキュメントを改めて整理してみると, TK はつぎのような特性をもつ知識であることがわかる. それは, 有体的な資源とアンバンドルしにくく, 公有であり, その保全と多様性の維持を求められ, かつ IP 的な価値ももち, 技術集積諸国の一方的な取得にさらされている.

このような TK に対して CBD の求めるような公正, 衡平な配分を実現できるかどうか.

4.2 項に示した課題について, とくに (1) 項について, 技術集積国と GR 保有国とが同意できるか否かにかかっている<sup>(3)</sup>.

### 参考文献

- [Dickson, David(1999)] "ICSU Seeks to Classify Traditional Knowledge", *Nature*, Vol.401, p.631.
- [Dutfield, Graham(2000)] "Intellectual Property Rights, Trade and Biodiversity", Earthscan Publications, passim.
- [Downes, David R.(2000)] "How Intellectual Property Could Be a Tool to Protect Traditional Knowledge", *Columbia Journal of Environmental Law*, Vol.25, pp.253-282.
- [Kelowna, Rex Dalton(2002)] "Tribes Query Motives of Knowledge Databases", *Nature*, Vol.419, p.866.
- [Khalil, Mohamed(1995)] "Biodiversity and the Conservation of Medical Plants", Timothy Swanson (ed.) "Intellectual Property Rights and Biodiversity Conservation", pp.232-253.
- [Schiermeier, Quirin(2002)] "Traditional Owners 'Should Be Paid'", *Nature*, Vol.419, p.423.
- [Shiva, Vandana(1997)] "Biopiracy" (松本文二訳 (2002) 『バイオパイラシー』, 緑風出版), passim.
- [UNESCO(1999)] "A Science Agenda".
- [WIPO(2001)] "Intellectual Property Needs and Expectations of Traditional Knowledge Holders", passim.
- [WIPO(2002a)] "Review of Existing Intellectual Property Protection of Traditional Knowledge", passim.
- [WIPO(2002b)] "Traditional Knowledge — Operational Terms and Definitions", passim.
- [大沢麻衣子 (2002)] 「知的財産としての伝統的知識の保護」, 『知財研フォーラム』, 50

<sup>(3)</sup>本解説は [名和, 2002c] の一部を大幅に書き換えたものである.

- 号, pp.29-32.
- [大塚善樹 (2002)] 「生物多様性から知的財産権の多様性へ」, 『現代思想』, 30 卷 11 号, pp.136-151.
- [高倉成男 (2001)] 『知的財産法制と国際政策』, 有斐閣, pp.339-356.
- [寺嶋秀明 (2002)] 「フィールドの科学としてのエスノ・サイエンス」, 寺嶋秀明他 (編) 『エスノ・サイエンス』, 京都大学学術出版会, pp.3-12.
- [名和小太郎 (2002a)] 『ゲノム情報はだれのものか』, 岩波書店, pp.95-160.
- [名和小太郎 (2002b)] 『学術情報と知的所有権』, 東京大学出版会, pp.157-161, 293-314.
- [名和小太郎 (2002c)] 「伝統的知識の知的所有権化」, 『科学技術社会論学会 — 第 1 回 年次研究大会予稿集』, pp.217-220.
- [山名美加 (2002)] 「知的財産権と先住民の知識」, 『現代思想』, 30 卷 11 号, pp.152-164.

(2002 年 12 月 17 日受付)

(2002 年 12 月 22 日採録)

## 長瀬真理先生の訃報に接して

本会理事，静岡大学情報学部教授，長瀬真理先生には，2002年8月7日，肺癌のため，享年55才にして逝去された。長瀬先生が，1988年の本会創立当初から，役員として本会活動に多大の貢献をなされてきたことは，会員の等しく謝するところである。ここに追悼の一文を綴って，些かなりとも先生のご功績に報いることとする次第である。

長瀬真理先生は，東京女子大学文理学部哲学専攻のご出身で，さらに東京都立大学大学院人文科学研究科に進まれ，ギリシア哲学を研究された。その後，プラトンの文体研究へのコンピュータの応用を契機として，コンピュータによる文章解析に関心を発展させ，この種のソフトウェアの始祖ともいべき OCP (Oxford Concordance Program) のわが国への導入に尽力された。この成果は「長瀬真理，・西村弘之著『コンピュータによる文章解析入門：OCP への招待』，オーム社，1986年」に表れている。随分以前のこととて，記憶が定かでないが，筆者が先生の知遇を得たのも，この前後ではなかったかと思う。筆者はこの頃，データベースの関連で東京大学大型計算機センターによく出入りしており，長瀬先生も OCP の同センターへの導入，利用に関して積極的に活動しておられたからである。

先生は，さらに日本文化の国際発信という見地から，日英仏対訳による源氏物語のハイパーテキスト化に取り組み，その成果は，“UNESCO Global Heritage Pavilion” 中の “The Tale of Genji in the Twenty-first Century” としてみるができる (<http://webworld.unesco.org/genji/index.shtml>)。こうして先生は国際的に活躍され，オックスフォード大学とシャルル・ド・ゴール・リール第3大学に客員教授として招かれ，さらに2002年秋にはハーバード大学において講義することになっていたのである。

さて，本会の活動に直接関連するご功績であるが，筆者のパソコンを探したところ，1988年12月19日の本会第1回理事会の議事メモが見つかった。そこには「企画・編集委員，長瀬真理氏より，人文・社会系部会におけるテキスト・データベース連絡委員会設立の提案があり云々」といった記述が見え，本会設立当初からの先生のご活躍が偲ばれる。また1991年からは，編集委員長として，論文誌，ニューズレターの編集に尽力されたことも記憶に鮮明である。藤原鎮男会長の主宰に係る1994年のFID大宮大会に際しては，英国からの有力研究者招請のため，ペリー英国図書館研究開発部長をロンドンに共に訪ねたことも改めて思い出される。また，筆者の属する学術情報センター（当時）とリール第3大学の交流に関連して，リール在任中の先生にご協力を仰いだこともあった。今，本会ニューズレター51号（1998.12.1）の巻頭言，「ボンジュール！」と題したリールからの帰朝報告を読めば，感慨を新たにするばかりである。その後も，2001年11月にしばらく入院とのメールを受け取るまで，学会活動，研究活動の各般に亘りご指導，ご協力を賜ったものである。

今般，突然の訃報に接し，只々驚き，痛恨の情はまさに筆舌に尽くし難い。ここに，夫君沖田好正氏により配信された，長瀬先生直筆の「自己死亡通知」の終節を引用し，切に先生のご冥福を祈りつつ，この追悼の文を結びたいと思う。

「ソクラテスが『弁明』の最後で言っているように、『しかし，もう終わりにしましょう。時刻ですからね。もう行かなければならないのです。わたしはこれから死ぬために，諸君はこれから生きるために。しかしわれわれの行く手に待っているものは，どちらがよいか，誰にもはっきりとはわからないのです。神でなければ。』

Adieu !!!」

2002年10月  
情報知識学会副会長 根岸正光  
(国立情報学研究所)



## 「情報知識学会」編集委員会規程

2002年8月27日  
制 定

### (設置)

第1条 情報知識学会定款第7章に基づき、本会に編集委員会（以下、「委員会」という）を置く。委員会に関する事項は、この規程に定めるところによる。

### (任務)

第2条 委員会は、会長の諮問に応じ、情報知識学会誌（以下、「会誌」という）の編集に関する事項について審議し、会誌の発行に当たる。

### (組織)

第3条 委員会に委員長を置く。

2 委員長は定款第6章による編集長が担当する。

3 委員長は情報知識学会正会員（個人会員）の中から若干名を委員として任命する。

### (委員会)

第4条 委員会は委員長が招集し、その議長となる。

2 委員会に必要に応じて役職委員を置くことができる。

3 委員長に事故あるときは、あらかじめ委員長が指名する委員がその職務を行う。

### (任期)

第5条 委員の任期は2年とし、再任を妨げない。

2 委員長が交代したときは前項に関わらず任期は終了する。

### (委員以外の者の出席)

第6条 委員長が必要と認めるときは、委員以外の者を委員会に出席させることができる。

### (小委員会)

第7条 委員会に具体的、専門的な事項を検討するために小委員会を置くことができる。

2 小委員会に関し必要な事項は、委員会が別に定める。

### (庶務)

第8条 委員会の庶務は情報知識学会事務局において処理する。

### (雑則)

第9条 この規程の定めるもののほか、委員会の運営などに関し必要な事項は、委員会において別に定める。

第10条 この規程の改定は理事会の議決を経て行う。

### 付則

1 この規程は2002年8月27日から施行する。

## 「情報知識学会誌」編集規程

2002年8月27日  
制 定

0. 情報知識学会定款第6章による機関誌「情報知識学会誌（以下、会誌という）」の編集に関する事項は、この規程の定めるところによる。
1. 原稿の提出と受付および保管
  - 1.1 本学会会員は本学会が発行する会誌に投稿することができる。ただし、その内容は著者の責任とする。
  - 1.2 原稿の書き方ならびに投稿の手続きは別に定める「投稿規定」「執筆要領」による。
  - 1.3 投稿時の原稿はすべて事務局に提出する。
  - 1.4 事務局は受けとった投稿原稿の受付年月日を記録し、保管する。
  - 1.5 編集委員会は会員または非会員への依頼によって原稿を集めることができる。
  - 1.6 討論を主な内容とする原稿のときは、編集委員会はそれを討論相手に見せ、なるべくその回答の原稿を求める。
2. 原稿の審査と採否
  - 2.1 編集委員会は原稿を審査し、掲載するかどうかを決める。
  - 2.2 編集委員会は、原則として原稿の査読を会員または非会員に依頼する。ただし、この場合も最終的な審査は編集委員会の責任において行うこととする。査読結果は原則として3ヶ月以内に著者に通知する。
  - 2.3 編集委員会は投稿規定と査読結果に基づいて、著者に修正を求めることがある。修正の求められた原稿は修正後すみやかに返送するものとするが、原稿が3ヶ月以上たつて編集委員会に返送されてきたときには原則として新規の投稿原稿として取り扱う。
  - 2.4 編集委員会が掲載適当と認めたとき（以下、これを「受理」とよぶ）、編集委員会は受理年月日を記録し、著者に通知する。
  - 2.5 編集委員会が掲載不適当と認めたとき、その理由を明らかにした文書を著者に送り、その了解を得たうえで原稿を返却する。
  - 2.6 掲載された原稿は原則として返却しない。ただし、図についてはあらかじめ申し出があれば、返却する。
3. 原稿の掲載
  - 3.1 投稿原稿の掲載順は原則として受理の順とする。ただし、編集上の都合により順序を変更することがある。
  - 3.2 学会費など、本会に納入すべきものを著しく滞納している会員の投稿原稿は、それが納入されるまで、掲載を延期することがある。
4. 校正
  - 4.1 初校正は原則として著者が行う。第2校正以降は編集委員会が行うが、必要に応じて著者に依頼することがある。
  - 4.2 著者は、校正した原稿を指定期日までに編集委員会に返送する。返送が著しく遅れた場合には編集委員会で校正するか、次号に回すこともある。

- 4.3 雑誌発行の日時が迫っているときなど、時間的な制約を受けるときには、著者による校正を略し、編集委員会が校正を行う場合がある。
- 4.4 校正のときに、著者が原稿と著しく異なるように書換えを行ったときは、その掲載を次号以降に延期することがある。

## 5. 別刷

- 5.1 別刷（抜刷）は著者の申し出により可能であるが、実費負担とする。

## 6. 会誌の内容

- 6.1 広い意味での情報知識学に関連し、またその発展に貢献するもの（情報／知識の収集、整理、蓄積、検索および各種解析、利用などに関するもの）とする。
  - 6.1.1 研究論文（Research Paper）：オリジナルな研究論文で、内容の主要な部分が学術論文として他に公表されていないもの。
  - 6.1.2 事例／調査報告（Report）：情報知識学に関連したシステムなどの開発、利用、調査に関するもの。資料も含む。
  - 6.1.3 解説／展望（Review）：情報知識に関連した特定分野の論文や学説などを総括、解説、紹介、あるいは技術動向などを展望したもの。技術、研究上の処理、解析方法などに関する解説。
  - 6.1.4 論談（Proposal Paper）：情報知識学に関連した新たな知見の表明、提案など。
  - 6.1.5 討論（Discussion）：本会誌に掲載された論文についての学術的な討論。
  - 6.1.6 研究速報（Notes）：技術、手法、新事実などの簡単な報告。
  - 6.1.7 講座（Lecture）：情報知識学の各分野に関する基礎理論、技術の適用などについて、テーマを定めて系統的に説明するもの。
- 6.2 学会記事（News）
  - 6.2.1 本会の事業、運営などに関する報告、記事、資料。
  - 6.2.2 ニュース、お知らせ。最近出版された単行本やモノグラフの紹介。
- 6.3 講演（Lecture）
  - 6.3.1 特別号などにおける講演資料。
- 6.4 その他
  - 6.4.1 上にあげたもののほか、編集委員会が適当と判断した事項。

## 7. 規約の改訂

- 7.1 本規約の改訂は、編集委員会の議を経て、理事会の承認を得なければならない。

## 8. 施行

- 8.1 本規約は 2002 年 8 月 27 日から施行する。



## 「情報知識学会誌」投稿規定

2002年8月27日  
制 定

0. 情報知識学会誌編集規程による本会機関誌「情報知識学会誌（以下、会誌という）」への投稿に関する事項は、この規定の定めるところによる。

### 1. 投稿資格

投稿者の少なくとも1人は本会員でなければならない。ただし、編集委員会による依頼原稿の場合にはこの限りではない。

### 2. 投稿原稿

2.1 広い意味での情報知識学に関連し、またその発展に貢献するもの（情報／知識の収集、整理、蓄積、検索および各種解析、利用などに関するもの）とする。刊行時において未発表の原著でなければならない。本会誌の記事の種類を以下に示す。

2.2 投稿者は会誌記事の種類を明記して投稿しなければならない。ただし、編集委員会で変更することがある。

- (1) 研究論文 (Research Paper)：オリジナルな研究論文で、内容の主要な部分が学術論文として他に公表されていないもの。
- (2) 事例／調査報告 (Report)：情報知識学に関連したシステムなどの開発、利用、調査に関するもの。資料も含む。
- (3) 解説／展望 (Review)：情報知識に関連した特定分野の論文や学説などを総括、解説、紹介、あるいは技術動向などを展望したもの。技術、研究上の処理、解析方法などに関する解説。
- (4) 論談 (Proposal Paper)：情報知識学に関連した新たな意見の表明、提案など。
- (5) 討論 (Discussion)：本会誌に掲載された論文についての学術的な討論。
- (6) 研究速報 (Notes)：技術、手法、新事実などの簡単な報告。
- (7) 講座 (Lecture)：情報知識学の各分野に関する基礎理論、技術の適用などについて、テーマを定めて系統的に説明するもの。
- (8) 学会記事 (News)：本会の事業、運営などの報告、記事、資料など。
- (9) ニュース、お知らせ (News)：ニュース、お知らせ。最近刊行された単行本やモノグラフの紹介。
- (10) 講演 (Lecture)：特別号などにおける講演資料。
- (11) その他：編集委員会が適当と判断したもの。

2.3 会誌記事の種類のうち、(1) から (6) までは査読を行う。その他については編集委員会で編集を行う。

### 3. 投稿原稿

#### 3.1 原稿の形式

##### (1) 投稿時の原稿

以下の A, B のいずれかの体裁 でプリントされたワープロ原稿（横書き）4部の提出とする。その他、執筆に関する詳細は「執筆要領」を参照のこと。

A 刷り上り原稿を想定したレイアウト（A4判，2段組，20字×46行×2段）。

図、表は希望の位置に配置すること。

B ベタ打ち原稿 (A4判, 40文字×40行).

図, 表は, 1枚ずつ別の用紙に印刷すること. ベタ打ち原稿右余白に図表の挿入位置を朱書きすること.

(2) 採択決定後の原稿

以下の C, D の両方の形式で記録された電子媒体1部の提出とする.

原稿の送付にあたってはフロッピーディスクなどの適当な電子媒体とする. 詳細は提出時に事務局に相談のこと.

C Microsoft Word, 一太郎, DVI, PDF などの代表的なフォーマット.

D 図, 表は十分な品質で印刷できる形式 (JPEG, GIF など).

### 3.2 原稿の制限

(1) 原稿の長さを原則として次のように制限する.

|                          |         |          |
|--------------------------|---------|----------|
| 研究論文, 事例/調査報告, 解説/展望, 論談 | : 刷り上がり | 20 ページ以内 |
| 討論, 研究速報, 講座             | : 刷り上がり | 6 ページ以内  |
| ニュース他                    | : 刷り上がり | 2 ページ以内  |

(2) 図原稿 (原図) の大きさは A3 判を越えないものとする.

(3) 原則として, 図版も含めてモノクロ印刷とする. ただし, カラーでなければならぬ図版を使用する場合は, 別途編集委員会と相談する. なお, カラーページやページを超過する分については, 印刷費を著者の全額負担とする.

(4) 使用言語は日本語または英語とする.

### 4. 原稿の採否

投稿原稿の採否は, 専門家による査読の後, 編集委員会において決定する. 不採択となった原稿は, 編集委員長より理由を付して通知する.

### 5. 査読のプロセス

学会員の中から編集委員会が指名した査読者2名によって査読を行う. 内容によっては, 編集委員会は著者に照会し, 原稿の修正を求めたうえで, 再査読を行うことがある.

### 6. 校正のプロセス

採択が決定した投稿原稿は, 掲載原稿として著者に校正を依頼する. 著者による校正は原則として1回とする. その際, 字句の修正以外は原則として認めない.

### 7. 別刷

別刷 (抜刷) は著者の実費負担とする. 希望部数を事務局に申し出ること.

### 8. 投稿の手続き

原稿投稿時には下記の書類を添え, 原稿送付先に郵送する.

#### 8.1 必要書類

(1) 最初の投稿時

a. 投稿原稿整理カード: ホームページからコピーして, 必要事項を記入し, 印刷したものを1部. 掲載原稿整理カードと兼ねるので, コピーを保存しておくこと.

b. 紙媒体の原稿 (図, 表を含む): 4部.

なお, 投稿者は著者校正用に原稿のコピーを保存しておくこと.

c. E-mail による連絡票

・ 標題 (和文, 英文), 著者名 (和文, ローマ字), 所属機関/住所 (和文, 英文),

要旨 (和文, 英文). キーワード (和文, 英文). 刷り上り予定ページ数  
・連絡先: 著者 1 名の連絡先 (氏名, 所属機関/部局, 所属機関住所, 電話番号, Fax  
番号, E-mail アドレス).

なお, 投稿後の連絡は主として E-mail で行う.

(2) 採択決定後の投稿

- a. 掲載原稿整理カード: 投稿時のカードに追加事項を記述し, 印刷したものを 1 部.
- b. 3.1(2) に指定した電子媒体: 1 部.
- c. 3.1(2) の C の印刷出力 (プリントアウト): 1 部.

8.2 原稿の送付先

〒 110-8560 東京都台東区台東 1-5-1 凸版印刷 (株) 内  
情報知識学会事務局  
電話: 03-3835-5692 Fax: 03-3837-0368  
E-mail: LDE01013@nifty.ne.jp

8.3 原稿の受付

事務局が原稿を受け取った日を受付日とする. 受付の確認を 1 週間以内に投稿者の連絡先に E-mail で通知する. 不備のある投稿原稿は返送し, 再提出するものとする.

9. 原稿提出期日

投稿は随時とする. ただし, 特集号などは除く.

10. 著作権

- 10.1 機関誌『情報知識学会誌』に掲載された論文 (電子版を含む) の著作権 (著作財産権, copyright) は情報知識学会に帰属する.
- 10.2 掲載論文は冊子による出版の他, 電子的に蓄積し, 本会が行う情報提供サービスなどを通じて公開する.

11. 規定の改訂

- 11.1 本規定の改訂は, 編集委員会の議を経て, 理事会の承認を得なければならない.

12. 施行

- 12.1 本規定は 2002 年 8 月 27 日より施行する.
- 12.2 本規定の施行により, 現行規定 (第 4 版 (暫定板) 2002 年 3 月) は廃止する.

## 「情報知識学会誌」執筆要領

2002年8月27日

制 定

### 1. 一般的な事項

本会誌への投稿は、「投稿規定」に従い、投稿原稿は本執筆要領に従って作成されなければならない。

本会誌の投稿原稿の種類には、研究論文、事例／調査報告、解説／展望、論談、討論、研究速報、講座、本会記事、講演、ニュース、その他がある。

### 2. 日本語原稿の構成

#### 2.1 全体構成

##### (1) 第1ページ（査読者には見せない）

- ・ 標題（和文および英文）
- ・ 著者名（和文およびローマ字，ローマ字による著者名は，名，姓の順で，姓は全て大文字を使用する。）
- ・ 所属（和文および英文による所属機関名）
- ・ 住所（和文による所属機関の住所．E-mail．脚注とする．）
- ・ 見出し用原稿（研究論文，事例／調査報告，解説／展望，論談の原稿には，刷り上がりページ上部欄外につける著者名および標題を30字以内で書く．）

##### (2) 第2ページ目以降（査読者に見せる）

- ・ 要旨（研究論文，事例／調査報告，解説／展望，論談の原稿には，和文および英文で要旨をつける．和文要旨の長さは400字以内とする．英文要旨の長さは500語以内とする．要旨中には，図，表，数式などを用いない．本文中の図，表，数式，文献などを番号で引用しない．）
- ・ キーワード（研究論文，事例／調査報告，解説／展望，論談，討論，研究速報，講座にはキーワードをつける．和文および英文でそれぞれ5個程度，和文と英文のキーワードは，対応することが望ましい．キーワードはカンマ（,）で区切る．）
- ・ 本文（和文または英文）
- ・ 文献，付録など（和文または英文）
- ・ その他（とくに長い論文の場合，読者の便宜を考えて内容目次を付してもよい．ただし，章，節の見出し程度とする．）

#### 2.2 本文（Body）

##### (1) 構成

章，節などの構成は，第1レベルは1，2，…，第2レベルは1.1，1.2，…，第3レベルは1.1.1，1.1.2，…のようにする。

##### (2) 脚注

脚注はできるだけ避ける．止む無く使用する場合は簡潔な文とする。

##### (3) 図および表

- a. 図，表にはそれぞれ通し番号をつける．図1 (Fig.1)，図2 (Fig. 2)，…表1 (Table 1)，表2 (Table 1)，…など。
- b. 通し番号とともに説明文（キャプション）をつける．キャプションの位置は図は上部に，表は下部とする。

- (4) 数式, 化学式
  - a. 数式 (独立式), 化学式は, 段落外で記述されているものも本文中で一回は参照する.
  - b. 数式には, 通し番号を振る.
- (5) リスト (または箇条書き)
  - a. 記号なしリスト.
  - b. 記号つきリスト. リストの記号は, 数字, アルファベット, 記号を用いることができる. ただし, これらの混在した使用は避ける. アルファベットは1 論文中では大文字, 小文字の使い分けをしない.
  - c. 複雑化を避け, せいぜい2 段 (親子関係) のリストとし, ネストを跨ぐ順序づけを用いない.
- (6) 注記および参考文献  
本文中で少なくとも一回は参照すること. 通し番号で参照し, タイトルなどでの参照は避ける.

### 2.3 後付け (End)

- (1) 謝辞  
本文の最後に続けて記述する. 章番号は用いない. 章題は「謝辞」とする. 最終原稿時に記述することが望ましい.
- (2) 注記および参考文献
  - a. 注記または参考文献には, 参照順に通し番号を付し, 本文の最後に番号順にまとめて記述する. 章番号は持たない. 章題は「参考文献」とする.
  - b. 1 つの番号には1 つの注記または参考文献を対応させる.
  - c. 注記中には参考文献を含めない. 注記はできる限り簡潔に表現すること.
  - d. 参考文献の記述形式は, 以下の形式を満たさなければならない.
  - e. URL を参照してもよいが, 移動または削除される可能性があるため, 極力避ける. 原著が URL でのみしか参照できない場合など, やむをえない場合は用いてもよい. その場合, 参照時点でのハードコピーを保管しておくなど, 参考文献へのアクセス手段を確保するよう努力しなければならない.

### 【参考文献の形式】

- 1. 雑誌中の1 論文  
[引用通し番号] 著者名: 論文名, 雑誌名, 巻号, 掲載ページ, 出版年, その他.
- 2. 図書1 冊  
[引用通し番号] 著者名: 書名, 版表示, 出版地, 出版社, 総ページ数, 出版年, その他.
- 3. 図書の1 部  
[引用通し番号] 著者名: 論文名, 書名, 版表示, 出版地, 出版社, 掲載ページ, 出版年, その他.
- 4. 会議報告  
[引用通し番号] 著者名: 論文名, 書名 (会議名), 版表示, 編集者名, 会議開催地, 会議開催年, 会議開催機関, 出版地, 出版社, 掲載ページ, 出版年, その他.
- 5. インターネット上の論文  
[引用通し番号] 著者名や標題など可能な限り詳細な書誌事項, URL, 参照年月日. (単なるホームページなどは参考文献にしないこと).

【参考文献の記述】

1. 著者名, 編集者名の記述

- (1) 個人著者名は, 姓, 名の順に記述する. 欧文著者名は, カンマ (,) で姓, 名を区切る.
- (2) 複数著者の場合は, 各著者をセミコロン (;) で区切る.
- (3) 翻訳図書などの翻訳者名の場合は, 著者名の後に括弧 ( ) に入れて記述する.

2. 論文名, 書名の記述

- (1) 論文名, 書名は, 和文の場合はかぎ括弧 (「」), 欧文の場合はダブルクォーテーション (“ ”) に入れて記述する.
- (2) 図書中の一部を引用した場合の書名は, 和書の場合は二重かぎ括弧 (『』) に入れ, 欧文の場合はイタリック体で記述する.

3. 掲載ページの記述

- (1) 論文の場合は, 開始ページと終了ページを記述する. 「pp. 開始ページ-終了ページ」とする.
- (2) 図書の場合は, 総ページ数とする. 「総ページ数 p.」とする.

【参考文献の記述例】

- [1] 藤原 謙: 「情報知識学試論」, 情報知識学会, Vol.1, No.1, pp.3-10, 1990.
- [2] 原正一郎; 安永尚志: 「国文学研究支援のための SGML/XML データシステム」, 情報知識学会, Vol.11, No.4, pp.17-35, 2002.
- [3] Fujiwara, Shizuo: “East-West Communication and Information Transfer — Coordination of Specificity”, Journal of Japan Society of Information and Knowledge, Vol.4, No.2, pp.11-18, 1994.
- [4] Ellis, David (細野公男監訳, 斎藤泰則, 鈴木志元, 村上泰子訳): 「情報検索論」, 丸善, 180p., 1994.
- [5] 根岸正光: 「学術情報の流通と利用」, 『情報学とは何か』情報学シリーズ 3, 丸善, pp.43-69, 2002.
- [6] 名和 小太郎: 「デジタル図書館と著作権」, デジタル図書館, No.4, [http://www.dl.ulis.ac.jp/DLjournal/No\\_4/nawa/nawa.html](http://www.dl.ulis.ac.jp/DLjournal/No_4/nawa/nawa.html) (2002年8月27日参照)

3. 文章と文体

- 3.1 文体はひらがなと漢字による口語常態 (である調) とし, 現代かなづかいを用いる.
- 3.2 漢字は当用漢字とする. ただし, 固有名詞や学界で広く用いられている慣用の術語はこの限りではない.
- 3.3 句読点その他には 「,」 「.」 を用いる.
- 3.4 本文中の人名には敬称をつけない. ただし, 謝辞の人名はこの限りではない.
- 3.5 数量を表す数字はアラビア数字とする.
- 3.6 数式は印刷に便利なように十分注意して記号を記すこと. 原則として数量 (変化量) を表す記号はイタリックとする.
- 3.7 ローマ字の人名の姓は大文字体とする.
- 3.8 固有名詞で読み誤るおそれのあるものにはふりがなをつける.
- 3.9 英数字は原則として半角英数文字で記述する.

4. 英文原稿

英文による投稿原稿の場合も、原則として和文による投稿原稿の諸規定に従う。英語圏以外の著者の場合、著者名表記にその国語による表記を認めるが、可能な限り英文表記とする。

4.1 研究論文、事例／調査報告、解説／展望、論談、討論、研究速報などの原稿は英文でもよい。

4.2 英文原稿は語学的に難点の少ないものであることを必要とし、著者の責任において完全を期する。

4.3 英文原稿には、英文による要旨 500 語程度、ならびに日本語による 400 字以内の要旨をつける。ただし、著者が日本語を理解できない場合は日本語要旨を省略できる。

5. その他

原稿は和文または英文によるものとする。文章は語学的に難点の少ないものであることとし、著者の責任において完全を期する。編集委員会は語学的校正を行わない。

6. 要領の改訂

6.1 本要領の改訂は、編集委員会の承認を得なければならない。

7. 施行

7.1 本規定は 2002 年 8 月 27 日より施行する。

様式1 / 様式2

## 「情報知識学会誌」投稿原稿整理カード / 掲載原稿整理カード

1. 論文種別
2. 標題 (和文)  
   標題 (英文)
3. 著者名 (和文, ローマ字)
4. 所属機関名 (和文, 英文)
5. キーワード (和文, 英文)
6. 連絡責任者 (1名)  
   氏名, 所属機関 / 部局, 同住所, 電話番号, Fax 番号, E-mail アドレス
7. 送付投稿原稿  
   テキスト部分の枚数  
   図の枚数  
   表の枚数  
   付録の枚数
8. 図の返却希望 (YES, NO)
9. カラー図の有無 (原則として認めていないが, カラーでなければならない場合など, 全額実費著者負担)
10. 投稿日  
\_\_\_\_\_
11. 登録番号
12. 受付日 (再受付日)
13. 受理日
14. 送付掲載原稿  
   フロッピーディスクなど 枚数  
   プレインテキストプリント 枚数 (刷り上がり見本, 図表なども貼り込んだもの)
15. 別刷り (抜刷り) の希望部数 (全額実費負担)

### 【投稿原稿整理カード】

1. 標題を「様式1 投稿原稿整理カード」とし, 1 から 10 項目 (11 項目以降は採択後) をもれなく A4 判横書き 2 枚程度に, ワードプロでお作りください.
2. 投稿原稿と一緒に送り下さい.
3. なお, 投稿時には「投稿規定」にある E-mail による連絡票もお忘れなく, お送り下さい.

### 【掲載原稿整理カード】

1. 標題を「様式2 掲載原稿整理カード」とし, 全項目をもれなく A4 判横書き 2 枚程度に, ワードプロでお作り下さい. 掲載原稿と一緒に送り下さい.



## 論文賞案内

2002年12月4日  
編集委員会委員長 安永尚志

論文賞が創設されます。情報知識学会に、待望の論文賞が新設されることになりました。会員の皆様、奮ってご投稿下さい。

### 論文賞の創設

#### 1. 目的

最近の目覚ましい情報技術の進展において、我が情報知識学会の役割はますますその重要度を高めるものとなり、また社会全般からも大きな期待を寄せられてきています。当学会が社会の要請に応え、積極的な活動を行うことは極めて重要と考えられます。

そのため、当学会誌に論文賞を創設し、社会の負託に応える必要があると考えました。下記のように、来年から実施する予定で準備を開始しました。

論文賞を設けることにより、情報知識学会の社会的評価の向上に資し、かつ学会誌のレベルアップをはかることに資することが期待されます。さらに、会員の論文投稿のインセンティブにも寄与できると考えています。

#### 2. 論文賞授与の実施方法

1年間を通じて、その年の会誌の中から選定した優秀な研究論文を総会にて表彰するものとします。当面は論文賞とし、賞品、賞金は出ませんが、表彰状を発行します。総会においてその栄誉を讃えることとします。

選定は編集委員会で行い、理事会で承認し、決定します。

#### 3. 2004年論文賞

会員への周知、また準備の都合もあり、2004年5月の総会を第1回の授賞式とします。ただし、初回でもあり候補論文は2002年度（第12巻）発表論文を含め、2003年（第13巻）の研究論文から選定することとします。

#### 4. その他

詳細は未定ですが、論文賞受賞論文については総会で論文の解説を含め発表していただくか、あるいはその後の展開などを含めて会誌に寄稿していただくことも考えています。

## 情報知識学会第11回(2003年度)研究報告会 発表論文募集について

実行委員長 国沢 隆 (東京理科大学理工学部)

情報知識学会では平成15年5月24日(土)に、前回と同じ学術総合センター(東京・一ツ橋)の会場で総会とともに研究報告会を開催する予定です。この研究報告会の発表論文を下記要領で募集いたしますので、学会員の皆様どうぞ奮ってご応募ください。

### 1. 募集分野

- (1) 情報知識の構造解析, モデル化, 可視化, 知識発見
- (2) 情報・知識の表現, 生産, 組織化, 検索, 提供
- (3) 電子出版, 電子図書館
- (4) マルチメディア, 電子ミュージアム
- (5) 用語, シソーラス
- (6) 知識情報の流通と知的所有権
- (7) 専門分野における品質管理, 基準化
- (8) その他情報知識学に関連する諸研究・開発

### 2. 応募方法

発表論文題目, 著者名(連名の場合, 登段発表者に○印), 所属, 論文概要(200字以内), 連絡代表者の氏名, 住所, 電話・FAX番号, 電子メールアドレスを明記の上, 下記宛になるべく電子メールにてお申し込みください。

〒278-8510 千葉県野田市山崎 2641

東京理科大学理工学部 国沢 隆

E-mail: kunisawa@rs.noda.tus.ac.jp FAX: 0471-23-9767

**応募期限：2003年3月31日 採択可否通知：2003年4月7日**

### 3. 論文執筆・発表について

- (1) 採択されたものについては論文執筆要領の詳細を別途お知らせいたします。発表論文はA4版で4ページを目処とお考えください。4ページまでは無料ですが、それを超えると有料(1ページ1,000円)になります。
- (2) 発表時間は、質疑応答を含めて30分を想定しています。なお、論文提出がないと発表はできません。
- (3) 登段発表者は当学会員に限ります。当日入会も可能です。

## 特集「科学技術データの活用」の論文公募 Vol.13, No.3 (2003年7月刊行予定)

情報知識学会誌編集委員 国沢隆 (kunisawa@rs.noda.tus.ac.jp)  
菅原秀明 (hsugawar@genes.nig.ac.jp)  
宇陀則彦 (uda@slis.tsukuba.ac.jp)

### 1. 主旨

ネットワーク技術やデータベース技術の発展に伴い、科学技術データを共有し、生物、物理、化学、地学などの研究に活用する動きが活発になってきました。科学技術データを用いた研究は独自の発展をとげ、新たな研究領域を形成しつつあります。例えば、生物学と情報学の融合領域であるバイオインフォマティクスにおいては、膨大なゲノムデータを蓄積するところから生命現象に関する新たな知識獲得までの過程で、新しい情報処理技術が生まれてきています。これまでは新しい技術をどうデータに適用するかという視点で研究が行われてきましたが、今はデータから新しい技術が生まれる時代になっています。そこで、本委員会では、「科学技術データの活用」というテーマで特集を組む時機にきたと判断しました。奮って多くの論文の投稿をお願いいたします。

### 2. 公募テーマ

- (ア) 生命現象データ、材料データ、気象データ、地理データなど科学技術データの蓄積、表現、組織化、管理、分析、利用
- (イ) 科学技術データを対象としたデータベース構築技術、システム構築技術、データマイニング技術、データ解析ツールなど
- (ウ) 科学技術データの共有に関する国内および国際ネットワーク協力事例
- (エ) 科学技術データ利用に関する倫理、法整備など

### 3. 投稿原稿および修正原稿の締め切りは以下のとおりです。

投稿原稿締め切り：2003年4月11日（金）

修正原稿締め切り：2003年5月14日（金）

### 4. 原稿送付先

〒110-8560

東京都台東区台東1-5-1 凸版印刷（株）内情報知識学会事務局

TEL：03(3835)5692 FAX：03(3837)0368 E-mail：LDE01013@nifty.ne.jp

情報知識関連新刊図書一覧 2002年9～12月期

- 図書館・表現の自由・サイバースペース ロバート・S・ペック著 日本図書館協会 2002・8  
 データベース 速水治夫・宮崎収兄著 オーム社 2002・9 2,500円 [2,000円]  
 情報セキュリティ技術大全 Ross Anderson 著 日経BP社 2002・9 5,800円  
 情報数学の世界2 — パラドックスの不思議 有沢誠著 朝倉書店 2002・9 2,500円  
 入門量子コンピュータ ゲナディ・P・ベルマン他著 パーソナルメディア 2002・9 2,800円  
 「e-ビジネス」の主治医 IT コーディネータ 和田英男・松原恭司郎著 プレジデント社 2002・9  
 暴走するインターネット 鈴木謙介著 イースト・プレス 2002・9 1,500円 [1,600円]  
 情報社会と情報倫理 梅本吉彦編著 丸善 情報教育シリーズ 2002・9 2,400円  
 読んだ本はどこへいったか 鶴見俊輔著 潮出版社 2002・9 1,800円  
 情報化社会対話集5 青田吉弘著 ラッセル社 2002・9 1,800円  
 IT 革命時代の経済と政府 長谷川啓之・谷口洋志著 文真堂 2002・9 2,500円  
 新サイバーネットワークのしくみ — 情報産業の新たな展開 浦山重郎著 技術評論社 2002・10  
 報道は欠陥商品と疑え 鳥越俊太郎著 ウェイツ 2002・10 750円 [2,880円]  
 シミュレーションの思想 広瀬通孝・小木哲朗著 東京大学出版会 2002・10 2,800円  
 電子社会のパラダイム 辻井重男・竹内啓他著 新世社 2002・10 2,380円  
 次代のIT戦略 高橋徹・永田守男編 日本経済評論社 2002・10 2,300円  
 インターネットで文献探索 2002年版 伊藤民雄著 実践女子大学図書館編集 日本図書館協  
 会 2002・10 1,800円  
 情報リテラシー 海野敏・田村恭久著 オーム社 2002・10 2,800円  
 情報技術と社会 中村維男編著 根元義章他著 東北大学出版会 2002・10 1,500円  
 一人から始まる — 情報社会のパーソナルデザイン 青田吉弘・加藤明弘著 ラッセル社 2002・11  
 情報科学 ヒューマン編 中易秀敏他著 共立出版 2002・11 2,900円 [1,800円]  
 ナレッジサイエンス 北陸先端科学技術大学院大学知識科学研究科監修 杉山公造編著 紀伊国  
 屋書店 2002・12 2,500円 [4,900円]  
 情報と自己組織化 ハーマン・ハーケン著 シュプリンガー・フェアラー東京 2002・12  
 知の挑戦 エドワード・O・ウィルソン著 角川書店 2002・12 2,200円  
 暗号技術のはなし ドリス・ベイカー著 ピアソン・エデュケーション 2002・12 3,800円

(注) 図書館流通センター刊『週刊新刊全点案内』を参照

コンピュータ関連の技術・操作関連のものは除外した。また、初心者入門的なもの、ビ  
 ジネス色の強いもので除外したものもある。

制作者 平田 周 2002/12/25

## 情報知識学会誌の発行変更について

情報知識学会誌編集委員会  
委員長 安永尚志

2003年より、情報知識学会誌の発行方式が変わります。従来は会計年度（4月から翌年の3月）内の通巻号として編集発行してきましたが、2003年より暦年（1月から12月）編集とすることになりました。したがって、2003年は第13巻として、1号を2月、2号を4月、3号（特集号）を7月、4号（特別号）を10月に発行する予定です。なお、2004年以降の1号は1月発行となります。

本誌（第12巻4号）は、以上の処置のため2002年12月に刊行を計画しましたが、諸般の事情で、通常通り2003年1月刊行としました。

---



---

情報知識学会誌 編集委員会

編集委員長 安永 尚志 国文学研究資料館  
副編集委員長 宇陀 則彦 筑波大学図書館情報学系

## 編集委員

|       |              |       |            |
|-------|--------------|-------|------------|
| 石塚英弘  | 筑波大学図書館情報学系  | 伊藤鉄也  | 国文学研究資料館   |
| 神立孝一  | 創価大学経済学部     | 国沢 隆  | 東京理科大学理工学部 |
| 阪口哲男  | 筑波大学図書館情報学系  | 菅原秀明  | 国立遺伝学研究所   |
| 中川 優  | 和歌山大学システム工学部 | 名和小太郎 | 国際大学       |
| 二階堂善弘 | 茨城大学人文学部     | 西脇二一  | 奈良大学社会学部   |
| 根岸正光  | 国立情報学研究所     | 原田隆史  | 慶應義塾大学文学部  |
| 藤原 譲  | 工業所有権総合情報館   | 細野公男  | 慶應義塾大学文学部  |

---



---

## 平成 15～16 年度役員選出について（中間報告）

前号（12 巻 3 号）でお知らせしたとおり，平成 15 年 4 月から 2 年間就任する役員の選出手続きを進めています。

第一段階として，正会員全員から役員候補者の推薦を募りましたが，平成 14 年 11 月末の締切日までに推薦は皆無でした。そのため，理事会は理事会として推薦する役員候補者リストを作成中です。今後の日程は以下のようになります。

- (1) 情報知識学会誌 13 巻 1 号に役員候補者リストを掲載し，正会員全員から投票を求める。投票用紙（返信葉書）は，総会出欠票および委任状を兼ねる。
- (2) 投票結果は，平成 15 年 5 月 24 日（土），学術総合センター（一ツ橋）の 2 階会議場にて開催する情報知識学会平成 15 年度総会で発表し，承認を得たうえ，役員を確定する。

以 上

情報知識学会事務局

### 年会費納入について

1. ご自分が納入した年月日の確認をしてください。

お手元に郵送された情報知識学会誌の封筒に貼ってある宛名ラベルをご覧ください。最下行に納入年（西暦下 2 桁）、月（2 桁）、日（2 桁）の計 6 桁が印字してあります。印字には、納入日から 10 日ほどかかります。

2. 「未納」と印字してある場合。

次のいずれかの方法で納入してください。1 年分の年会費は、正会員 8 千円、学生会員 4 千円です。

- 郵便振込口座 00150-8-706543 情報知識学会（代表 藤原鎮男）
- UFJ 銀行 秋葉原駅前支店 普通預金 3586133 情報知識学会（会長 藤原鎮男）

所属団体名で支払う場合、郵便振込票の通信欄に個人名を明記してください。銀行振込では個人名が記載されませんので、電子メール・FAX・葉書など、別ルートで団体名と会員名（個人名）をお知らせ願います。

3. 年会費の納入期限

毎年 5 月末までに納入してください。会計年度は 4 月 1 日より翌年 3 月末までです。退会するかたは新年度に入る前（3 月末まで）に電子メール、FAX、葉書などの文書で、退会届をご提出ください。その際、年会費の滞納分はお支払い頂きます。

■複写をされる方に

本誌に掲載された著作物を複写したい方は、(社)日本複写権センターと包括複写許諾契約を締結されている企業の従業員以外は、著作権者から複写権等の行使の委託を受けている次の団体から許諾を受けて下さい。著作物の転載、翻訳のような複写以外の許諾は、直接本会へご連絡ください。

〒107-0052 東京都港区赤坂 9-6-41 乃木坂ビル 学術著作権協会  
TEL: 03-3475-5618 FAX: 03-3475-5619 E-mail: naka-atsu@muji.biglobe.ne.jp

アメリカ合衆国における複写については、次に連絡してください。

Copyright Clearance Center, Inc. 222 Rosewood Drive, Danvers, MA. 01923, USA  
TEL: 978-750-8400 FAX: 978-750-4744 URL: <http://www.copyright.com/>

情報知識学会誌 Vol.12, No.4 2003年1月28日発行 編集・発行 情報知識学会  
頒布価格 3000円

---

情報知識学会 (JSIK: Japan Society of Information and Knowledge)

会長 藤原 鎮男

事務局

〒110-8560 東京都台東区台東 1-5-1 凸版印刷(株)内  
TEL: 03(3835)5692 FAX: 03(3837)0368 E-mail: LDE01013@nifty.ne.jp  
URL: <http://www.jsik.jp/>

---

# Journal of Japan Society of Information and Knowledge

## Contents

Special Issue: XML based system for information and knowledge

### Foreword

Foreword ..... Hidehiro ISHIZUKA ... 1

### Research Paper

An Application of EAD/XML to the Japanese Archival Descriptions and the Management of Archives: The Position of EAD, Encoded Archival Description in the Management Process of Archives ..... Haruyoshi GOTOH ... 3

An XML Format for Proteomics Database to Accelerate Collaboration among Bioresearchers ..... Hironori MIZUGUCHI, Ken'ichi KAMIJO, Akira TSUGITA ... 22

Electronic Data Interchange Protocol for Healthcare  
..... Shoichiro HARA, Hiroki SUGIMORI, Katsuhiko FURUMI,  
Ikuo TOFUKUJI, Takeshi KUBODERA, Masaki KAWAI, Katsumi YOSHIDA ... 32

Research and Development of a Web Retrieval System on the Concept of "Computer-Processable Structured Digital Object based on XML"  
..... Ying LI, Hidehiro ISHIZUKA ... 53

### Proposal Paper

A Community Aims at "Knowledge-Emergent Society" by "Document Management"  
..... Ken NISHIMURA ... 69

### Review

Traditional Knowledge and Intellectual Property Rights: A Review  
..... Kotaro NAWA ... 77

### Memorial

Obituary, Professor Mari NAGASE ..... Masamitsu NEGISHI ... 84

### Information

Information for Authors ..... Hisashi YASUNAGA ... 85

Establishment of JSIK Award ..... Hisashi YASUNAGA ... 96

Call for Papers on Conference ..... Takashi KUNISAWA ... 97

Call for Papers on Special Issue  
..... Takashi KUNISAWA, Hideaki SUGAWARA, Norihiko UDA ... 98

Books ..... Shu HIRATA ... 99

情報知識学会誌 第12巻4号 2003年1月28日発行

編集兼発行人 情報知識学会 〒110-8560 東京都台東区台東1-5-1 凸版印刷(株)内

TEL: 03(3835)5692 FAX: 03(3837)0368 (振替: 00150-8-706543)

学術刊行物 ISSN 0917-1436