

*Journal of Japan Society of Information and Knowledge*

# 情報知識学会誌

Vol.13 No.1 (Mar. 2003)

~~~~~ 目次 ~~~~~

|      |                                                                 |    |
|------|-----------------------------------------------------------------|----|
| 論文   | 生物学の概念ネットワークの特性について<br>..... 真栄城 哲也、藤原 譲、下原 勝憲                  | 1  |
| 論文   | データの再利用が可能な複数用語体系ブラウジングシステムの構築<br>..... 高久 雅生、江草 由佳、伊東 千夏、石塚 英弘 | 10 |
| 論談   | 文化財情報の記述項目 — 現状の分析と系統化の課題 —                                     | 23 |
| 調査報告 | 「日本古典文学本文データベース」の利用調査                                           | 33 |
| 報告   | 2002年情報知識学会月例懇話会報告                                              | 42 |
| 再掲   | 投稿規定と執筆要領                                                       | 47 |
| お知らせ | 論文賞創設                                                           | 55 |
|      | 第11回研究報告会 発表論文募集                                                | 56 |
|      | 特集「科学技術データの活用」の論文公募<br>..... 国沢 隆、菅原 秀明、宇陀 則彦                   | 57 |



情報知識学会

# 生物学の概念ネットワークの特性について

## On Properties of Biological Concept Network

真栄城 哲也<sup>†</sup> 藤原 譲<sup>‡</sup> 下原 勝憲<sup>†</sup>

Tetsuya MAESHIRO Yuzuru FUJIWARA and  
Katsunori SHIMOHARA

生物学の専門書、辞典や用語データベースから 133,251 語の専門用語を抽出し、それらの専門用語間の意味関係を自動的に抽出した。抽出された専門用語は生物学の概念を表し、専門用語間の意味関係は概念間の意味関係を表す。生物学の概念をノードで表し、概念間に意味関係があればリンクで結ぶことによって得られる生物学概念ネットワークの特性を解析した。生物学概念ネットワークは、ネットワークの規模とは独立した特性であるスケールフリー構造を 2 重に持つ。また、生物学での概念の重要性と、概念毎に意味関係で結ばれる概念の数は、少なくとも重要性が上位の概念では比例している。これら上位の重要性もしくは意味関係の数を持つ概念は、概念の検索時間を短くする役割を担っている。

133,251 terms extracted from biological books and dictionaries were processed to automatically extract semantic relations among terms such as hierarchical and equivalent relations. Extracted terms represent biological concepts, thus extracted relations define the biological conceptual network. Analysis of the constructed conceptual network shows double scale-free property, and concepts with large number of semantic relations are shown to be important biological concepts. These concepts with large number of semantic relations shorten the time to search concepts in the network.

キーワード：概念ネットワーク、生物学、意味関係、スケールフリー  
Conceptual network, Biology, Semantic relationship, Scale free

### 1 はじめに

生物学の概念をノードで表し、概念間に意味関係があればリンクで結ぶことによって得られる生物学概念ネットワークの特性を解析した。生物学概念ネットワークの特性は、生物学の知識全体の特性を概念と意味関係の側面から捉えたものであり、このような特性が明らかになれば、バイオインフォマティクスに代表される生物学の情報処理、表現、記録に最適な手法に関する手掛かりが得られる。

ネットワーク構造の解析は、Small World Network(SWN)<sup>[1]</sup>の研究によって様々な自然

界のネットワークや人工ネットワークが解析されている<sup>[2]</sup>。今回の解析によって生物学概念ネットワークは、ネットワークの規模とは独立した特性であるスケールフリー構造(巾分布構造)を 2 重に持つことが判明したが、このような 2 重スケールフリー構造のネットワークは殆んど知られていない。

本論文では、Small World Network の特徴と、生物学の概念および意味関係の自動抽出方法について述べた後、構築した生物学概念ネットワークの特性について考察する。

### 2 Small World Network

ネットワークの構成要素をノード、そしてノード間の関連関係をリンクで表現する。ノード間の接続が何らかの規則性に従ってい

<sup>†</sup> ATR 人間情報科学研究所

ATR Human Information Science Laboratories  
maeshiro@atr.co.jp

<sup>‡</sup> 工業所有権総合情報館

National Center for Industrial Property Information

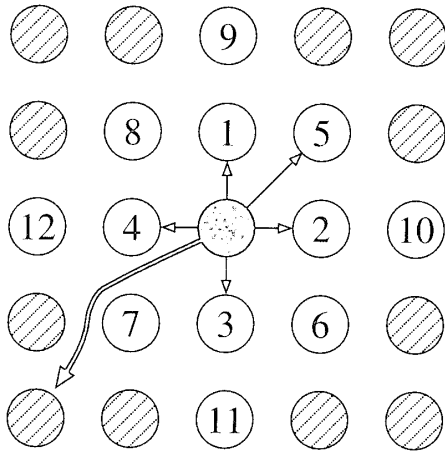


図1 任意なランダムさのネットワークの生成. ○は対象のノードであり,  $k$  の値によって, 番号付けられたノード (近傍ノード) に順次接続する. 確率  $p$  で, 近傍ノード以外のノードからランダムに選択されたノードに接続する.

るネットワークを規則的なネットワーク, そして規則性の無い接続のネットワークをランダムなネットワークとする.

ネットワークのランダムさを  $p$  ( $0.0 \leq p \leq 1.0$ ) で表す.  $p = 0.0$  は規則的,  $p = 1.0$  はランダムなネットワークを意味する. 例えば,  $n$  個のノードがあり, それぞれのノードに平均  $k$  個の他のノードが接続されているネットワークを考える. 単純なネットワークの生成例として, 2次元平面上のネットワークでノード間の接続に  $k$  と  $p$  の2つのパラメータを用いる (図1). 任意のランダムさ  $p$  のネットワークを以下のように生成する. 確率  $p$  ( $0.0 \leq p \leq 1.0$ ) でランダムな接続を生成し,  $1.0 - p$  の確率で近傍のノードに接続する.  $k$  は各ノードに接続するノード数であり, ランダム確率  $p$  は,  $k$  個の内, 近傍ではなく遠距離のノードに接続する確率である. この例ではランダム接続の相手先ノードは等確率で選ばれる. 従って, ノード当たり平均で  $kp$  個の接続がランダムに他のノードと接続される.  $p = 0.0$  の場合, 規則的なネットワークが生成され ( $kp = 0.0$ ),  $p = 1.0$  の場合にはランダムなネットワークが生成される ( $kp = k$ ).  $p$  の増加とともにランダムに接続されるリンクが増え, ランダム

なネットワークに近くなる.  $p$  を変動させることで, 規則的なネットワークからランダムなネットワークへと順次生成可能である.

このようにして生成したネットワークの構造の尺度として, 全ノード間の最短距離の平均である平均最短距離  $L$  と, ノードがどれだけ密集しているかを表すクラスタ係数  $C$  を用いる [1].  $C$  は各ノードの近傍に存在する全ノード間に存在するリンクの密度の平均である. ノード  $i$  に  $k_i$  個のノードが接続しており, それら  $k$  個のノード同士間のリンク数が  $m$  である場合, そのノードのクラスタ係数  $C_i$  は,

$$C_i = m \left/ \frac{k_i(k_i - 1)}{2} \right. \quad (2.1)$$

で与えられる. なお,  $k_i(k_i - 1)/2$  は  $k_i$  個のノードが完全結合した場合のリンク数である. また, ネットワークのノード数を  $n$  とするとクラスタ係数は

$$C = \frac{1}{n} \sum_{i=1}^n C_i \quad (2.2)$$

で与えられる.

$L$  は大局的なパラメータであり,  $C$  は局所的なパラメータである. 従って,  $L$  が小さい程あるノードからネットワーク全体への情報伝達は速く,  $C$  が大きい程ノードは近傍のノードと密に接続しており, 局所的な情報伝達が速い. 前者はある概念に順次関連する概念を検索できる速度と関連し, 後者は, 局所的な情報の共有度および局所的な概念の関連密度とも関連する.

$L$  と  $C$  の値は, 規則的なネットワーク ( $p = 0.0$ ) では大きく,  $p$  の増加とともに減少し, ランダムなネットワーク ( $p = 1.0$ ) では最小となるが, 値の変動は非線形であり, 連動もしない [1].

Small World Network (SWN) [1] は, 局所的には規則的なネットワークの特徴を持ち, そして大局的にはランダムなネットワークの特徴を持つ, 両者の中間に位置するネットワークである. 線虫の神経細胞の回路網, 単細胞生物の代謝経路, インターネットのホームペー

ジのリンクや電力分配網など、実世界の様々なネットワークが SWN である [2].

SWN は、特に  $p$  が小さい場合のネットワーク構造を指し、ランダムなネットワークのように平均最短距離  $L$  が小さく、規則的なネットワークのようにクラスタ係数  $C$  が大きい、両者の特徴を持ったネットワークである。

なお、ネットワークのノード数を  $n$ 、ノード当りのリンク数の平均を  $k$  とした場合、

$$n \gg k \gg \ln(n) \gg 1 \quad (2.3)$$

の条件を満たすランダムなネットワークにおいて、全ノードが結合する条件は

$$k \gg \ln(n) \quad (2.4)$$

である [3]. また、 $k$  が  $\ln(n)$  に近くなるほど、あるノードからネットワーク全体への情報伝達速度は、規則的なネットワークやランダムネットワークよりも SWN が高速である [4].

SWN が規則的なネットワークやランダムなネットワークと明らかに異なるもう 1 つの特徴は、ノード当りのリンク数  $k$  の分布である。規則的なネットワークは、図 1 から判るように、 $k$  は定数である。一方、ランダムネットワークでは  $k$  はポアソン分布である。SWN の場合、 $k$  は巾分布が多く、そうでない場合でも対数分布等、明らかに定数およびポアソン分布とは異なる分布を持つ。 $k$  が巾分布であることは、 $k$  がネットワークの大きさ (ノード数) に依存しないスケールフリーなパラメータであることを意味する。

SWN の平均最短距離  $L$  の小ささは、少数のランダムな接続によるノード間の距離の短縮が理由だが、別の観点からも説明できる。ノード当りのリンク数が巾分布であることは、リンク数が大きなノードが少数存在することを意味する。従って、リンク数が大きなノードが頂上に位置し、下へ行くほどリンク数が少ないノードを配置する木構造に似たネットワークの表現手法を用いれば、ノード間の移動には、まずリンク数が多いノード (ハブノード) へ移動し、そこから目的のノードへ移動すれば距離が短くできることが判る。このよう

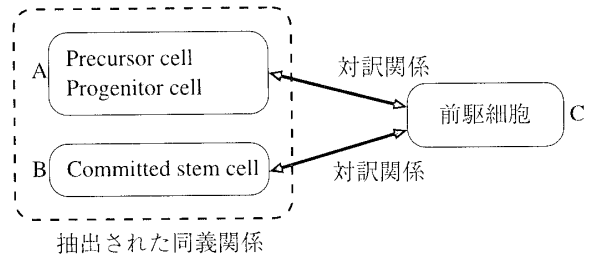


図 2 C-TRAN によって抽出された意味関係の例。左側 (ボックス A と B) は英語、右側 (ボックス C) はそれに対する日本語である。英語の 2 つのボックス (A と B) 内の専門用語は英語の情報源からそれぞれ独立に抽出された同義関係を表しており、英語の情報源のみを用いた場合はこれらは互いに無関係である。それぞれの日本語訳を用いることで、ボックス A とボックス B の専門用語は同じ日本語の用語「前駆細胞」に対応することが判明する。その結果、ボックス A とボックス B の計 3 つの専門用語は、ボックス C の日本語の専門用語を介して同義であることが抽出される (点線ボックス)。

に、ハブノードによっても短い平均最短距離を説明できる。

### 3 概念と意味関係の抽出

#### 3.1 概念の抽出

概念間の意味関係を抽出するには、概念が必要である。今回対象としている生物学のような自然科学の場合には、専門用語が対象分野の概念を表していると考えられる。専門用語を MeSH\*の見出し語、分子生物学、脳神経学、生化学、免疫学、微生物学等の分野の 26 冊の専門書 (付録参照) の索引に記載の見出し語、そして辞典の見出し語から抽出した。英語と日本語の 2 ケ国語の専門用語を扱い、対訳関係も抽出した。このようにして抽出した専門用語は、重複を除いて 133,251 語あり、これらを基に意味関係を抽出した。

\* <http://www.nlm.nih.gov/mesh/>

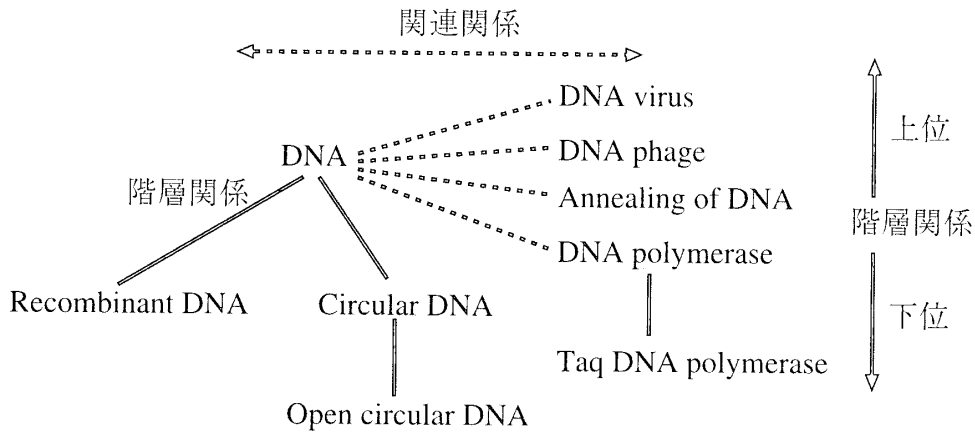


図3 SS-KWEICによって抽出された意味関係の例。実線は階層関係、点線は関連関係を表す。修飾子が前に付くと、専門用語の意味を限定するために階層構造で下位概念、そして後に付くと関連関係であると見なす。例えば、“circular DNA”は“DNA”の下位概念であり、“DNA”と“DNA virus”は関連関係である。また、“Annealing of DNA”のように、前に付いても“of”が存在するため関連関係と見なす(DNA annealingと同義)。

| 6-ホスホフルクトキナーゼ |                       |
|---------------|-----------------------|
| 対訳            | 6-phosphofructokinase |
| 同義            | ホスホヘキソキナーゼ            |
| 関連            | フルクトース・ビスホスファターゼ      |
| 関連            | 解糖系                   |
| 関連            | 調節酵素                  |
| 関連            | ホスホエノールピルビン酸          |
| 関連            | 無益回路                  |

図4 辞典から抽出された意味関係の例。見出し語と、その用語の説明内容に記載されている見出し語と関連する用語。抽出された用語は、見出し語に関連する用語として辞典の説明文中に記号が付けられているもの。

### 3.2 概念間の意味関係の抽出

専門用語(概念)の収集より重要なのは、概念間を関係付ける意味関係の抽出である。概念間の関係に基づいて概念を構造化して初めて、収集した専門用語の活用が可能となる。

抽出した概念間の意味関係の種類は、同義関係、関連関係、類似関係、階層関係、属性、である。意味関係の抽出は専門用語の生成規則に基づいたC-TRAN(図2)とSS-KWEIC(図3)<sup>[5][6]</sup>、辞典に記載の内容(図4)、そしてMeSHに記載されている概念の階層構造も利

用した(図5)。また、専門用語の抽出に利用した26冊の専門書(付録)の本文を用いて意味関係を抽出した。専門書の本文から意味関係を抽出する手法として、既に抽出済の専門用語が同じ文に共出現すれば関連関係で結んだ。

このようにして2,574,099個の意味関係を抽出した。意味関係抽出の処理で、全体の99.5%である132,563個の概念について意味関係を見出し、概念ネットワークを構築した。1つの概念当りの意味関係の平均数(ノード当りの平均リンク数)は19.3個である。

## 4 考察および結論

概念毎の意味関係の数 $k$ のヒストグラムを取ると、1,000個前後の意味関係を持つ概念を境に2重のスケールフリー構造を持つことが解る(図6)。この構造は、意味関係数(リンク数) $k$ の増加とともにヒストグラムの値が指数的に減少する構造とは異なる。

図6は、得られた概念ネットワークがSmall World Network(SWN)と類似の構造を持っていることを示している。このことは、概念間の距離を、概念を接続する意味関係の最短数と定義した場合、多くの概念と意味関係を持つハブとして機能する少数の概念の存在(表1)

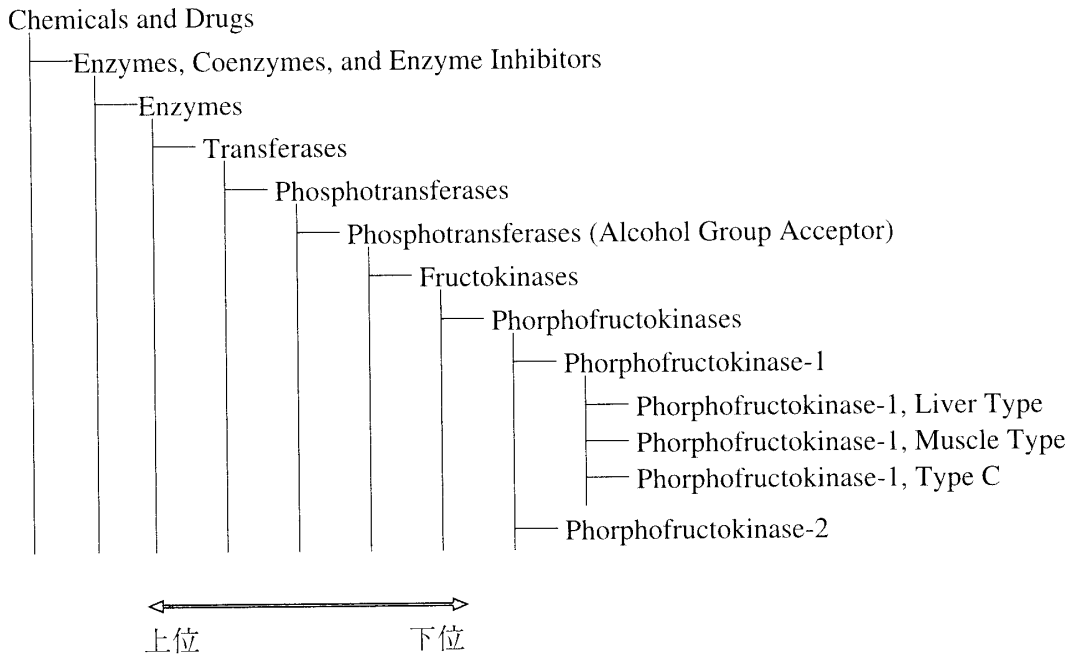


図5 MeSHに記載されている階層構造の例。木構造ではなく、複数の親(上位概念)を持つ概念も存在する。

表1 ハブ概念として機能する上位20個の専門用語

|            |                  |
|------------|------------------|
| 1 Cell     | 11 Kinase        |
| 2 Protein  | 12 Antibody      |
| 3 Gene     | 13 Peptide       |
| 4 Virus    | 14 Mutation      |
| 5 Disease  | 15 Tissue        |
| 6 Enzyme   | 16 Animal        |
| 7 Membrane | 17 Human         |
| 8 Antigen  | 18 Transcription |
| 9 Growth   | 19 Hormone       |
| 10 Mouse   | 20 Tumor         |

によって、ある概念から他の概念へ短距離で到達できることを意味する。これらのハブ概念(表1)は、生物学の知識体系で重要な概念であり、生物学の専門家ならば必ず知らなければならない概念である。これは、SWNの特徴であるハブノードの構造である。また、本研究で抽出した意味関係を用いた場合、生物学の重要な概念は他の重要な概念を直接連想させると考えられる。このことは、特定の概

念の重要性がネットワーク構造から推測できることを示唆している。このように意味関係の数によって概念の重要性を定量化する本手法は、従来よく使われている文章内の専門用語(概念)の出現頻度を用いる手法とは根本的に異なる。

意味関係数  $k$  から捉えた概念ネットワークの構造が巾分布構造を持ち、スケールフリー構造であることは、概念ネットワークの形成および成長がランダムではないことを示唆している。ネットワークの規模が新しいノードの追加によって拡大していくネットワーク成長過程について計算機シミュレーション実験を行い、新しいノードから既存のノードへの接続と、既存のノード間の新しい接続の計2種類の接続を追加する複数のアルゴリズムを比較した結果、少なくともランダムな接続では巾分布構造を持つネットワークは生じないことが判っている。図6が示唆するのは、概念ネットワークにおいては、新しい概念は意味関係を多く持つハブ概念に選択的に接続しやすいということである。従って、新しい概念は重要な概念に関連付けられて概念ネット

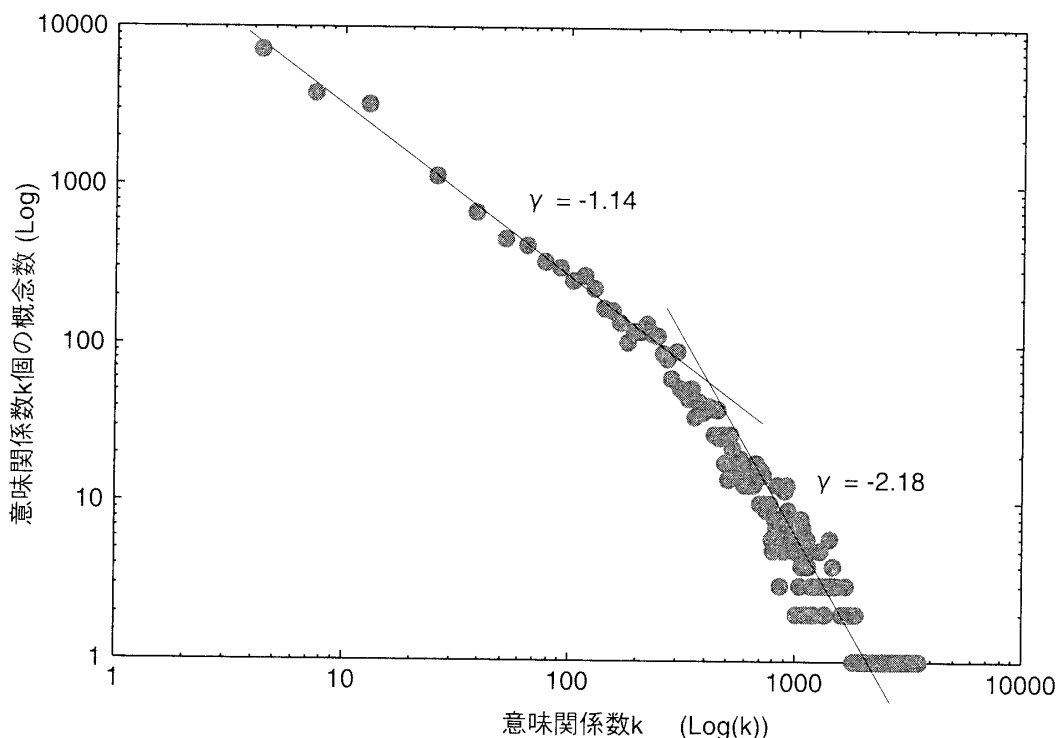


図6 概念毎の意味関係数の分布.  $k$  個の意味関係を持つ概念数  $P(k)$  は  $k^\gamma$  に比例する. 左部分は  $\gamma = -1.14$ , 右部分は  $\gamma = -2.18$ .

ワークに組み込まれる.

巾分布が見られる現象に Zipf の法則があるが, 本論文で扱っている意味関係から構築した概念ネットワークの巾分布とは異なる. 意味関係の数の分布は, Zipf の単出現や共出現の頻度とは異なる現象である. なお, 2重のスケールフリー構造が, 生物学の概念構造特有か, 概念構造に共通するかの解析が必要である.

SWN の1つの特徴は, ノード間の平均最短距離  $L$  が短いことである. 生物学の概念ネットワークのノード間の平均最短距離は  $L = 2.31$  と短い. この値は意味関係に基づいて接続された概念間の平均最短距離に相当し, 概念間の距離が短いことを示している. このことは, 概念間の探索が少ないステップ数で可能であり, 検索も高速に実行できることを意味する. また, ある概念に関連する概念の抽出も少ないステップ数で可能である.

さらに, 概念ネットワークの平均最短距離  $L$  は概念数の増加の影響をほとんど受けない特徴を持つ. 平均最短距離  $L$  は理論的に

$$L = \frac{\log(n/k_1)}{\log(k_2/k_1)} + 1 \quad (4.1)$$

で与えられる [2]. なお,  $n$  は概念ネットワークの概念数,  $k_1$  は概念に直接つながっている概念数 (1ステップで到達できる概念数), そして  $k_2$  は2ステップで到達できる概念数である. 従って, 概念数が10倍 (概念数約130万) になってもたかだか  $L = 2.7$ , さらに100倍 (概念数約1300万) でもたかだか  $L = 3.0$  である. これらの値は最悪値であり,  $k_1, k_2$  ともに増加しない場合の平均最短距離の値であるが, 実際はどちらの値も増加するため, 実際のノード間の平均最短距離は理論値よりも小さいことが予想される.

一方, もう1つの構造パラメータであるクラスタ係数の値は  $C = 0.14$  である. グラフ理論から, ランダムネットワーク ( $p = 1.0$ ) のクラスタ係数は

$$C_{\text{rand}} = \frac{k}{n} \quad (4.2)$$

で与えられる. なお,  $k$  はノード当りの結合数の平均,  $n$  はノード数である. 概念ネット

ワークと同じノード数とノード当りの結合数の場合、 $C_{\text{rand}} = 10^{-4}$  であり、SWN の条件であるランダムネットワークよりも大きなクラスタ係数 ( $C = 0.14 \gg C_{\text{rand}} = 10^{-4}$ ) を満たしている。

概念ネットワークにおけるハブ概念の役割も含めて、概念ネットワークにおける概念間の接続特性の解析として、概念がいくつの意味関係を持つ概念と接続しているかを解析できる (図 7)。意味関係数  $k$  が与えられた時、接続している概念の意味関係数が、対象としている概念の持つ意味関係数  $k$  より多いか少ないかで分類し、対象の概念と同じクラス ( $k$  以下) に接続している割合を計算する (この値を  $K_L$  とする)。図 7 は、意味関係数を  $k$  個持つ概念に接続している  $k$  個の概念の内、 $k$  個より多い意味関係数を持つ概念の割合である (この値を  $K_G$  とする。  $K_G = 1.0 - K_L$ )。

概念毎の意味関係数の分布 (図 6) とは異なり、意味関係数 2,000 個前後以上のハブノードの領域に巾分布構造が見られる。また、図 7 の曲線はランダム性を意味するポアソン分布とは明らかに異なる。

一方、意味関係数が少ない概念においても、構造に興味深い特性が見られる (図 7)。図 7 の点線は概念ネットワーク全体で意味関係数  $k$  より多い意味関係を持つ概念の割合を重み付きで計算した値であり、以下の式で与えられる。

$$K_{G\text{theory}}(k) = 1.0 - \frac{1}{N_w} \sum_{i=1}^k n(i) \times i \quad (4.3)$$

$$N_w = \sum_i^{k_{\text{max}}} n(i) \times i \quad (4.4)$$

なお、 $n(i)$  は  $i$  個の意味関係を持つ概念の数、 $k_{\text{max}}$  は対象ネットワークの概念が持つ意味関係の最大数、そして  $N_w$  は概念数の重み付き総和値である。この点線 ( $K_{G\text{theory}}(k)$ ) との比較の理由は、SWN の形成方法として概念間 (ノード間) の接続確率が意味関係数 (リンク数) に比例する接続形態が知られているからである。この方法でネットワークを生成す

る場合、 $K_{G\text{theory}}(k)$  は意味関係数を  $k$  個持つ概念が意味関係数  $k$  より多い概念と接続する確率を表す。

図 7 の 2 つの曲線の比較から次の推測ができる。概念ネットワークの形成が、SWN の 1 形成方法である接続確率が意味関係数 (リンク数) に比例した接続によるならば両曲線は同じはずであるが、図 7 では異なっている。  $K_G$  の値は単純減少せず、 $k = 10$  以降で一旦増加する。意味関係数を 1 つしか持たない概念の殆んどが意味関係を 2 つ以上持つ概念と接続しているのは当然である。しかし、意味関係数  $k$  が 2 から 15、特に  $2 \leq k \leq 10$  の範囲で同じか少ない意味関係数を持つ概念と継っている割合が今回構築した概念ネットワークの特徴である。図 7 の値  $K_G$  は意味関係数が  $k$  より多い概念と継っている割合であるため、 $1.0 - K_G$  が意味関係数  $k$  以下の概念と継っている割合である。従って、 $2 \leq k \leq 15$  の範囲の意味関係数を持つ概念は小規模なクラスタを形成しており、 $k = 20$  前後を境にした階層構造を持つと考えられる。

より詳細なネットワーク特性の解析には、より詳細な意味関係の抽出が必要である。同じ概念間には複数の意味関係があり、概念構造を捉える観点に応じて有効な意味関係が異なってくる。意味関係の抽出処理で問題になるのは、誤りの検出である。概念間の意味関係を自動抽出する場合、誤った意味関係がほぼ間違いなく混入するため、自動生成された概念構造に含まれる誤った意味関係の自動検出の研究が必要である。

さらには、ある概念に注目した場合、その概念を中心とするネットワークは観点によって異なる。例えば、単純な分類の場合でも、ある概念の属性数が  $a$  の場合、その捉え方は  $2^a$  通りあり、観点の数が  $2^a$  通りあることに相当する。さらに、分類方法は  $a!$  通りある。一方、因果関係や階層関係の上位・下位概念等、意味関係によってはリンクに指向性を持たせる必要がある。従って、ネットワーク特性の解析に用いた平均最短距離  $L$  およびクラスタ係



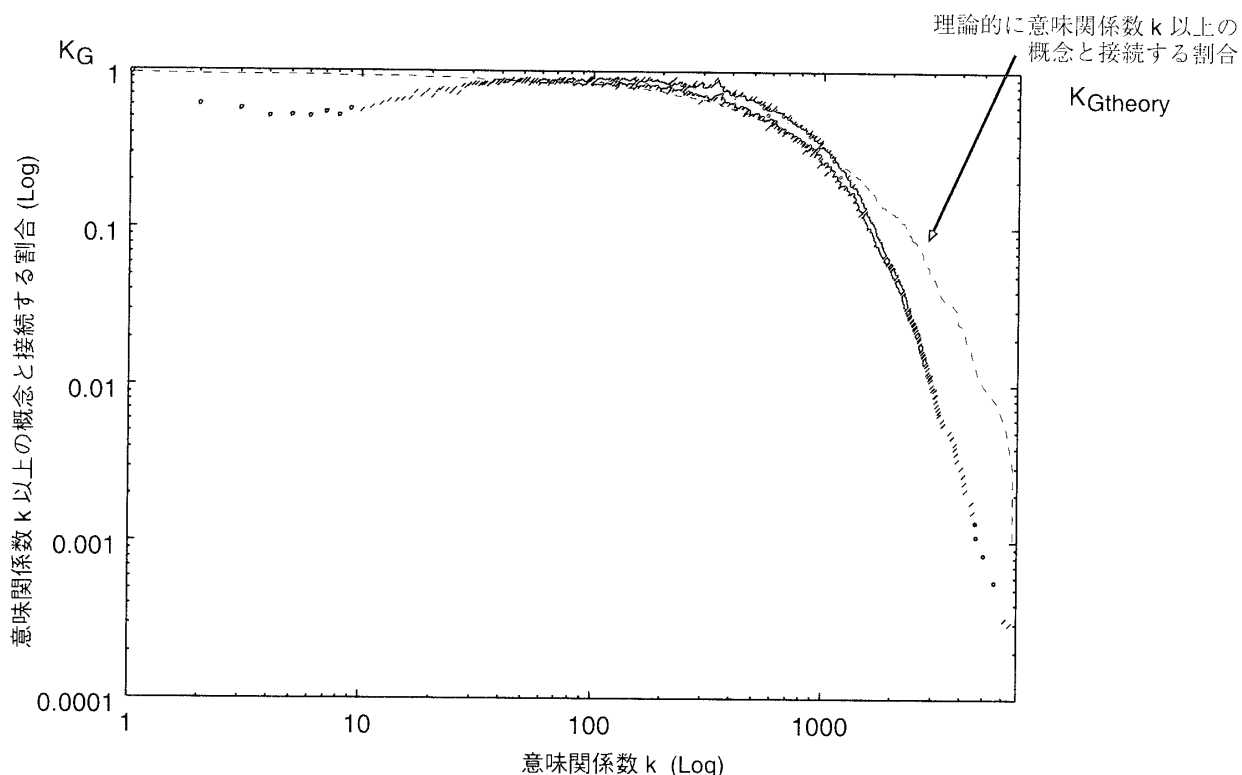


図7 概念の接続と意味関係数の関係. ある概念と接続している概念について, 意味関係数  $k$  より多い意味関係を持つ割合 ( $K_G$ ). 点線は, 概念ネットワーク全体で意味関係数  $k$  より多い意味関係を持つ概念の割合を式 (4.3) と (4.4) を用いて計算したもの.

数  $C$  を観点に依存する定義に変更し, 計測する必要がある.

#### 謝辞

本研究は通信・放送機構の研究委託により実施したものである.

#### 参考文献

[1] Watts, D.J.; Strogatz, S.H.: "Collective dynamics of 'small-world' networks", Nature, Vol.393, pp.440-42, 1998.  
 [2] Albert, R.; Barabasi, A.-L.: "Statistical mechanics of complex networks", Reviews of Modern Physics, Vol.74, pp.47-97, 2002.  
 [3] Ballabás, B.: "Random Graphs", Academic Press, 1985.  
 [4] Maeshiro, T.; Ohi, N.; Shimohara, K.: "Analysis of small-world networks with

high broadcast speed", Proc. of the 6th International Symposium on Artificial Life and Robotics, pp.27-30, 2001.

[5] Fujiwara, Y.; Lai, J.: "An information-base system based on the self-organization of concepts represented by terms", Terminology, Vol.3, pp.313-334, 1997.  
 [6] Fujiwara, Y.; Lee, W.G.; Ishikawa, Y.; Nishioka, A.; Hatada, H.; Fujiwara, S.: "Dynamic thesaurus for intelligent access to research databases", Proc. 47 FID Conf. (Helsinki), Vol., pp.173-181, 1988.

#### 付録

専門用語と意味関係を抽出した文献のリストである.

- 今堀, 山川監修:「生化学辞典 第3版」, 東京化学同人, 1998.
- 大沢, 小山, 奥田, 矢田編:「免疫学辞典」, 東京化学同人, 1993.
- Cruse; Lewis: “Atlas of Immunology”, CRC Press, 1998.
- Voet; Voet: “Biochemistry”, John Wiley, 1995.
- Michal: “Biochemical Pathways”, John Wiley, 1999.
- Riddle; Blumenthal; Meyer; Priess: “C. elegans II”, Cold Spring Harbor Lab. Press, 1997.
- Abbas; Lichtman; Pober: “Cellular and Molecular Immunology”, W.B.Saunders, 2000.
- Lodish; Berk; Zipursky; Matsudaira; Baltimore; Darnell: “Molecular Cell Biology”, Freeman.
- Wilson; Hunt: “Molecular Biology of the Cell Problems Book”, Garland.
- Lackie; Dow: “Dictionary of Molecular and Cellular Biology”, Academic Press, 1999.
- Zigmond; Bloom; Landis; Roberts; Squire: “Fundamental Neuroscience”, Academic Press, 1999.
- Paul: “Fundamental Immunology”, Lippincott-Raven, 1999.
- Fields; Knipe; Howley: “Fundamental Virology”, Lippincott Williams & Wilkins, 1995.
- Lewin: “Genes VII”, Oxford Univ. Press, 2000.
- Goldsby; Kindt; Osborne: “Kuby Immunology”, W.H. Freeman, 2000.
- Stedmans: “Medical Dictionary”, Lippincott Williams & Wilkins.
- Scheffler: “Mitochondria”, Wiley-Liss, 1999.
- Levitar; Kaczmarek: “Neuron”, Oxford Univ. Press, 1997.
- Strand: “Neuropeptides”, MIT Press, 1999.
- Wolpert: “Principles of Development”, Oxford Univ. Press, 1999.
- Kandel; Schwartz; Jessel: “Principles of Neural Science”, McGraw-Hill, 2000.
- Flint; Enquist; Krug; Racaniello; Skalka: “Principles of Virology”, ASM Press, 2000.
- Krauss: “Biochemistry of Signal Transduction”, Wiley-VCH Verlag, 1999.
- Lehninger: “Principles of Biochemistry”, Freeman, 1995.
- Siegel; Agranoff; Albers; Fisher; Uhler: “Basic Neurochemistry”, Lippincott Raven, 1998.
- Alberts; Bray; Johnson; Lewis; Raff; Roberts; Walter: “Essential Cell Biology”, Garland, 1997.
- Silverman: “Organic Chemistry of Enzymes”, Academic Press, 2000.
- Ferscht: “Structural Mechanisms of Proteins”, Freeman, 1999.

(2002年7月9日初回受付)

(2003年1月7日再受付)

(2003年1月21日採録)

データの再利用が可能な  
複数用語体系ブラウジングシステムの構築  
A Browsing System Based on Reusable Data Format  
for Various Terminological Systems

高久 雅生<sup>†</sup> 江草 由佳<sup>†</sup> 伊東 千夏<sup>†,1</sup> 石塚 英弘<sup>†,2</sup>

Masao TAKAKU Yuka EGUSA

Chinatsu ITO and Hidehiro ISHIZUKA

本研究では、用語体系データの再利用による用語体系の作成支援を目的として、複数用語体系のブラウジングシステムを構築した。本システムでは、用語間の階層関係をクリックابلマップ形式、HTML形式などとして図示することにより、利用者はその構造を直観的に把握できる。また、複数の用語体系の持つ用語とその階層関係を重ね合わせて表示することにより、利用者はそれらの構造を把握し、作成している体系に適した語彙を収集できる。また、システムが扱う用語体系データをXML形式で提供するため、データの再利用が可能となる。さらに、分野や規模の異なる5つの用語体系にこれらの手法を適用し、その有効性を示した。本システムの利用者は、用語体系の構造をブラウジングしながら、複数の用語体系を任意の組み合わせで重ね合わせた結果を閲覧したり、XML形式として取得して再利用したりすることができる。

We developed a browsing system for various terminological systems. Our system can support to create a new terminological system, using terms with hierarchical relationship picked up from existent terminological systems. A user can grasp the hierarchical relationships between selected terms intuitively, since our system illustrates the structure in clickable map form, HTML form, etc. Our system also integrates the terms with hierarchical relationships selected from several terminological systems, and displays the terms and relationships among them, therefore a user can grasp the structure and can collect the vocabularies suitable for the new terminological system for him/her to create. Moreover, our system provides reusable terminological data written in XML (Extensible Markup Language). Our system is successfully applied to five terminological systems which differ from each other in their fields or scales. Looking at the structure of several terminological systems, a user of our system can browse the integrated result of one or more terminological systems in arbitrary combination, and can also acquire the displayed data as the XML form, and can reuse it for his/her task. These results indicate effectiveness and usefulness of our system.

キーワード：用語体系, ブラウジングシステム, 重ね合わせ, データの再利用  
Terminological system, Browsing system, Integration, Reuse of data

<sup>†</sup> 図書館情報大学 情報メディア研究科  
Graduate School of Information and Media Studies,  
University of Library and Information Science  
masao@ulis.ac.jp

<sup>1</sup> 現在, 会社員

Presently a company employee

<sup>2</sup> 現在の本務は, 筑波大学 図書館情報学系

Presently with Institute of Library and Information  
Science, University of Tsukuba

## 1 はじめに

概念間の関係や用語の意味を記述した用語体系は、情報検索や自然言語処理などの分野で幅広く用いられてきた。近年では、ソーラス、分類表、Web Directoryなどのさまざまな用途で用語体系が作成されている。WWW(World Wide Web)上では、全分野

を網羅する約 44 万件からなる ODP(Open Directory Project)<sup>[1]</sup>, 生物や化学, 行政など幅広い分野から構成される環境用語を対象とした EIC ネット<sup>[2]</sup>における用語集など, 多種多様な用語体系が存在する. また, WWW 上の様々なサービスが独自の用語体系を提供することによって, 利用者の利便をはかろうとしている. また, 次世代の WWW とも言われる Semantic Web<sup>[3][4]</sup>でも, その基盤となるオントロジーの構築の重要性がうたわれている. このように用語体系は, 情報提供システムに欠かせない存在になりつつある.

これらの用語体系を作成し維持していくには, 以下のようなプロセスを踏む必要がある. まず, 体系の設計者は体系全体の設計を行い, その語彙を収集し, 分析を行なう. そして, それらの用語間の関係を一貫性をもって把握し, 管理しなければならない. しかし, 用語体系は一般に大量の用語と用語間の関係を含むため, これらのプロセスには多くのコストがかかっている.

筆者らは, 上記のような用語体系の作成, 維持を支援するには, 用語体系のブラウジングと複数用語体系の重ね合わせおよび再利用が有効であると考えた. これらの利点を以下に挙げる.

- ブラウジング

多くの用語とそれら同士の関係が記述された用語体系の構造を把握し分析するためには, 体系を文字として記述したものを眺めるよりも, これらの構造を概観できるように, わかりやすく可視化し, ブラウジングできるようにすることが有効である.

- 複数の用語体系の重ね合わせ

用語体系の作成の際には, 複数の用語体系がもっている用語とそれらの関係を重ね合わせて表示することが, 体系の把握および語彙の収集に有効である. 本稿では, これを重ね合わせと呼び, 複数の用語体系がもっている用語とそれらの関係を重ね合わせることを指す語として用いる.

- 用語体系の再利用

用語体系の重ね合わせを行うことにより, 似たような目的を持った既存の用語体系の持つ構造, 全く別の目的を持った体系の構造, 特定の分野の体系の構造などの一部を再利用することができ, それが有効な場合もある. 例えば, 網羅的な体系を作成する際に, 特定分野の体系を参照し, その構造を再利用したり, 逆に特定分野の体系を作成する際には, 汎用的な体系の構造を参照し, その構造を周縁領域の構造に再利用したりする場合である.

そこで, 本論文では上記で述べた, 用語体系のブラウジング, 重ね合わせ, 再利用が可能な用語体系ブラウジングシステムの構築について報告する. 本システムでは, 任意の見出し語とそれに関係する語を複数の体系間で自由に重ね合わせ, その結果をブラウジングできるようにした. このような重ね合わせの結果を元にして利用者(体系の設計者)が用語の取捨選択を行うことが, 複数用語体系の再利用に有効であると考えられる. さらに, システムが扱う複数の用語体系とその重ね合わせ結果を XML(Extensible Markup Language)形式<sup>[5]</sup>で提供することにより, 複数用語体系間での再利用やデータ交換が可能になるものと考えられる.

## 2 用語体系の重ね合わせ

本研究では, 用語体系間の再利用を行う際に用語体系を重ね合わせて, その結果をブラウジングすることが有効であると考え, その機能を設けた.

別々の目的で作られた用語体系間では, 体系が持っている見出し語に関する情報の量や詳細度, さらに, 何を見出し語とするかについてのとらえ方が大きく異なる. このため, それらの体系同士を重ね合わせるためには, それぞれの体系に共通の情報, つまり, どの体系も共通に持っている「見出し語」と「上位・下位関係」の情報のみを基礎として重ね合わ

せを行う必要がある。

筆者らはこれまで、それぞれの用語体系が共通に持っている情報だけを用いて重ね合わせを行う手法について実験を行ってきた<sup>[6]</sup>。この実験手法を元にして、本研究では、どの体系にも共通して存在する見出し語とその語の持つ階層関係を対象として、見出し語の照合のみにより用語体系の重ね合わせを行う手法を採用した。なお、本システムの目的は、利用者が重ね合わせ結果をブラウジングしながら、用語の取捨選択を判断するための判断材料を与えることにあるため、本手法による重ね合わせ結果の妥当性について、システムは判断しないこととした。

本研究では、用語体系を、見出し語をノードとし、見出し語間の階層関係をエッジとする有向グラフ構造としてとらえる。以下では、「上位語→下位語」として示す。

用語体系の重ね合わせは、複数の用語体系の部分グラフ同士を、同一の見出し語からなるノードを同一のノードとみなして結合し、各ノードの持つ関係、つまりエッジについてはそのまま保持することで行った。ここで、用語 A, B, C, D, E について階層関係が記述されている体系 1, 体系 2 がある場合 (図 1) を例にとって説明する。まず、用語 A および D が両体系に共通しているためこれらを結合する。さらに、用語 A の持つ階層関係 A → B (体系 1), A → D (体系 1, 2), A → C (体系 2), A → E (体系 2) をそのまま保持して、重ね合わせ結果とする (図 2)。

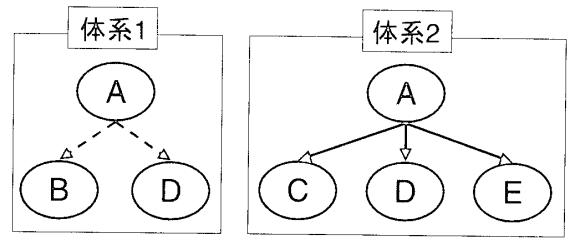


図 1 用語体系における階層関係の記述

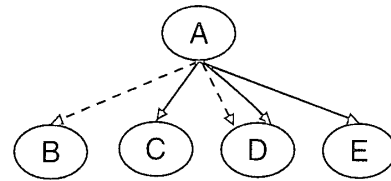


図 2 重ね合わせ例: 重ね合わせ結果

### 3 対象データ

本システムが対象とした用語体系は、以下の 5 つである。

- (1) EDR 専門用語単語辞書 (情報処理) 中の名詞・概念体系辞書 (以下 EDR)<sup>[7]</sup>
- (2) 日本十進分類表 新訂 8 版機械可読データファイル (以下 NDC)<sup>[8]</sup>
- (3) 国立環境研究所による環境情報ポータルサイト EIC ネット<sup>[2]</sup>上の環境用語集 (以下 EIC)
- (4) Open Directory Project<sup>[1]</sup>による Web ディレクトリの階層構造データ (以下 ODP)
- (5) ODP のデータのうち、「Kids and Teens」階層に限ったもの (以下 ODP-kt)

これらの用語体系が持つ用語数、階層関係数、その特徴を表 1 に示す。

表 1 各用語体系の用語数および階層関係数

| 用語体系   | 用語数     | 異なり語数   | 階層関係数   | 特徴                                      |
|--------|---------|---------|---------|-----------------------------------------|
| EDR    | 117,242 | 117,242 | 184,958 | 情報処理分野の専門用語を対象とした大規模な体系                 |
| NDC    | 30,659  | 25,410  | 102,639 | 全学問分野を対象とした大規模な体系                       |
| EIC    | 391     | 361     | 380     | 生物・化学・気象などの複合領域から構成される環境分野を対象とした体系      |
| ODP    | 446,926 | 150,387 | 446,928 | WWW 上に存在する情報全てを対象とした大規模な体系              |
| ODP-kt | 2,059   | 1,617   | 2,058   | WWW 上に存在する情報のうち、子ども向けの情報に限定した部分を対象とした体系 |

```

...
<Topic r:id="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット">
  <catid>499398</catid>
  <d:Title>インターネット</d:Title>
  <narrow r:resource="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/メーカー"/>
  <narrow r:resource="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/ブラウザ"/>
  <narrow r:resource="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/FTP"/>
  <narrow r:resource="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/ホームページ作成・
管理"/>
  <narrow r:resource="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション"/>
</Topic>
<Topic r:id="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション">
  <catid>510586</catid>
  <d:Title>コミュニケーション</d:Title>
  <narrow r:resource="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション/Odigo"/>
  <narrow r:resource="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション/AOL インスタント・メッセンジャー"/>
  <narrow r:resource="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション/ICQ"/>
  <narrow r:resource="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション/MSN_Messenger"/>
</Topic>
<Topic r:id="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション/Odigo">
  <catid>550024</catid>
  <d:Title>Odigo</d:Title>
</Topic>
<Topic r:id="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション/AOL インスタント・メッセンジャー">
  <catid>550027</catid>
  <d:Title>AOL インスタント・メッセンジャー</d:Title>
</Topic>
<Topic r:id="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション/ICQ">
  <catid>550138</catid>
  <d:Title>ICQ</d:Title>
</Topic>
<Topic r:id="Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション/MSN_Messenger">
  <catid>550286</catid>
  <d:Title>MSN_Messenger</d:Title>
</Topic>
...

```

図3 対象とした用語体系のデータ例：ODP

実際の用語体系のデータの一例として、用語体系 ODP における「Top/World/Japanese/コンピュータ/ソフトウェア/インターネット/コミュニケーション」階層を中心とした階層構造の XML データの抜粋を図3に示す。これは ODP プロジェクトが作成、配布しているデータ形式である。この例では、<Topic>要素に一階層についての情報が記述されており、<narrow>要素にはその階層の持つ下位関係が記述されている。また、<d:Title>要素にそ

の語の見出し語が記述され、<catid>要素にはその見出し語を一意に識別する ID が記述されている。<narrow>要素の r:resource 属性で示した ID が<Topic>要素の r:id 属性を示すことで、上位・下位関係の対応も知ることができる。

表1の用語体系を選択した理由は、様々な分野や規模の用語体系に本システムが適用できることを示すためである。本システムは、図3に示したような見出し語と上位・下位関

係が記述されたデータであれば対象としうる。

## 4 ブラウジングシステム

### 4.1 概要

本システムは、WWW 上からのユーザの要求に応じて動的に用語とその用語の持つ階層関係を可視化し、提示する。本システムは、各用語体系から作成されるインデックス、可視化・表示プログラムの2つから構成される(図4参照)。可視化の表示形式としては、グラフ構造のクリックブルマップ表示やHTMLによる一覧表示がある。さらに、表示された見出し語はリンクとなっており、次々とブラウザ上で各用語の持つ関係を閲覧していくことが可能である。これらの表示は、各用語体系および複数の用語体系の重ね合わせ結果のどちらにも適用可能である。

また、各用語とそれらの持つ関係を効果的にブラウジングするためには、対象データ全てを一度に表示するよりも、任意の見出し語を中心として、一定の範囲に表示を限定する必要があると考えた。そこで、利用者が指定した見出し語と、その語の上位語・下位語から数階層たどれる範囲までの用語を表示することとした。このうち、上位語・下位語についてそれぞれ、どの範囲の階層まで表示するかを利用者が指定できるようにした。

また、複数用語体系の重ね合わせに関しても重ね合わせたい体系の組み合わせをユーザが自由に指定できるようにした。例えば、 $n$ 個の体系がある場合には、

$$\sum_{i=1}^n n C_i$$

通りの組み合わせを指定できる。

また、各用語体系もしくは重ね合わせ結果を再利用するためのXMLデータを取得することもできる。

また、表示したい見出し語を選択するための機能には、見出し語検索(完全一致、部分一致)と最上位階層の見出し語一覧を表示する

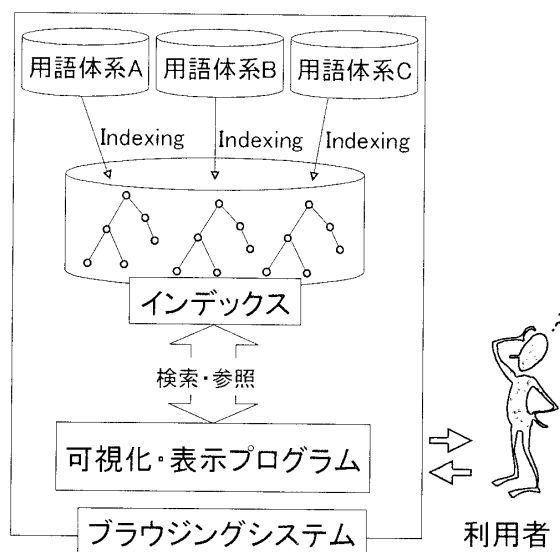


図4 システムの構成

機能がある。

なお、ユーザが本システムを利用するためには、Webブラウザさえあればよく、その他の特別なソフトウェアは必要としない。

### 4.2 インデックスの作成

本システムでは、複数の用語体系間での重ね合わせの組み合わせを任意に指定できるようにした。そのため、体系間の重ね合わせについては、結果をあらかじめ保持しておくのではなく、重ね合わせの対象となる用語体系の組み合わせを指定したブラウジングを利用者が要求するたびに動的に行うこととした。この動的な重ね合わせ処理を高速化するために、1)階層関係インデックス、2)見出し語検索用インデックスの2つのインデックスを作成した。

階層関係インデックスは、見出し語IDをキーとし、1)見出し語、2)上位語IDの集合、3)下位語IDの集合をそれぞれ値に持つ3つのPerlのハッシュファイルとして作成した。図3のデータに対して作成されるハッシュ表現を図5に例示する。

見出し語検索用インデックスは、見出し語を対象としたSuffix Arrayインデックスとし

```

%WORDS    = ("510586" => 'コミュニケーション');
%CHILDREN = ("510586" =>
              '550024|550027|550138|550286');
%PARENTS  = ("510586" => '510586');
    
```

図5 階層関係インデックスのPerlでの表現例

で作成した。

Intel PentiumIII 550 MHz×4, メモリ 2 GB のマシンでインデックスの作成に要する時間 (10 回の試行の平均) をそれぞれ以下に示す。

- EDR: 358 秒
- NDC: 33 秒
- EIC: 0.4 秒
- ODP: 954 秒
- ODP-kt: 2.1 秒

どの用語体系においても、十分に実用的な時間内でインデックスが作成できる。

これらのインデックスを作成することによって、利用者からのブラウジングのリクエストがあるたびに、高々1秒程度の実時間で、任意の見出し語の持つ階層関係を動的に生成することが可能になった。

### 4.3 可視化・表示プログラム

可視化・表示プログラムにおけるデータの流れを図6に示す。可視化を行う CGI プログラムでは、まず、4.2 節で述べたインデックスを参照して動的にある語を表わす XML データを生成する。次に、利用者が要求した表示形式にしたがって、その XML を元に XSLT スタイルシート<sup>[9]</sup>による変換を行い、提供するようにした。また、データの再利用のために、XSLT による変換をせず、XML データそのものを取得することもできる。

図7に用語体系 NDC の見出し語「図書館教育」について動的に生成した XML 表現を示す。XML 表現では、一用語の情報を <word> 要素として表現し、id 属性にその見出し語を示す ID

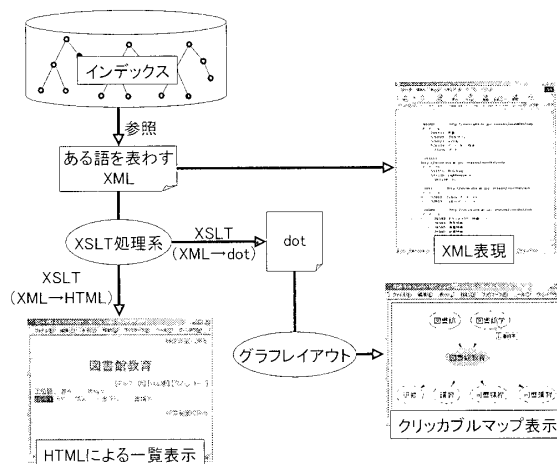


図6 可視化・表示プログラムのデータの流れ

```

<?xml version="1.0"?>
<wordlist>
  <word id="134">
    <name>図書館教育</name>
    <child idref="135">研修</child>
    <child idref="136">講習</child>
    <child idref="137">司書課程</child>
    <child idref="138">司書講習</child>
    <parent idref="127">図書館</parent>
    <parent idref="128">図書館学</parent>
  </word>
</wordlist>
    
```

図7 動的に生成した XML(NDC): 「図書館教育」

を記録し、<word>要素中の<name>要素で見出し語を表現した。<word>要素中の<parent>、<child>要素では、その見出し語のそれぞれ上位語、下位語を表現し、idref 属性にその上位語・下位語の ID を表現した。

さらに、図8に EDR, NDC, ODP の3つの用語体系に存在する見出し語「アニメーション」について重ね合わせた結果として動的に生成した XML を示す。XML の構造自体は図7の一体系での XML 表現とほぼ同一であるが、重ね合わせ結果における XML 表現では、<word> 要素に origin 属性を加えて、その見出し語がどの用語体系に由来するものであるかを表現した。

表示形式に HTML を指定した場合は、この



```

<?xml version="1.0"?>
<wordlist>
<word id="86468" origin="http://www.ulis.ac.jp/~masao/wordlist/odp">
  <name>アニメーション</name>
  <child idref="288444">声優</child>
  <child idref="576022">プロダクション</child>
  <child idref="576024">タイトル</child>
  <child idref="576319">アニメーター・作家</child>
  <parent idref="77596">アート</parent>
</word>
<word id="593554" origin="http://www.ulis.ac.jp/~masao/wordlist/odp">
  <name>アニメーション</name>
  <child idref="593556">POV Ray</child>
  <child idref="594116">Lightwave 3D</child>
  <parent idref="584979">3D</parent>
</word>
<word id="4894" origin="http://www.ulis.ac.jp/~masao/wordlist/edr">
  <name>アニメーション</name>
  <child idref="92012">フィルム・アニメーション</child>
  <child idref="92013">フィルムアニメーション</child>
</word>
<word id="28589" origin="http://www.ulis.ac.jp/~masao/wordlist/ndc">
  <name>アニメーション</name>
  <parent idref="28583">ドキュメンタリー映画</parent>
  <parent idref="28584">科学映画</parent>
  <parent idref="28585">各種映画</parent>
  <parent idref="28586">記録映画</parent>
  <parent idref="28587">教育映画</parent>
</word>
</wordlist>

```

図 8 動的に生成した XML(重ね合わせ結果): 「アニメーション」

XML に対して HTML 生成用の XSLT スタイルシートで変換した結果を返す。また、表示形式にクリッカブルマップや GIF 画像を指定した場合は、クリッカブルマップ生成用の XSLT スタイルシートで、グラフ表現の dot 形式<sup>[10]</sup>に変換し、その結果を dot コマンドでレイアウトした結果を返す。図 9 に XML 表現を dot 形式に変換する XSLT スタイルシートの例を示し、図 10 には Graphviz<sup>[10]</sup>でグラフィックレイアウトされた画像例を示す。

なお、動的に生成する XML の形式は、表示形式には依存していないため、新たな表示形式に容易に対応可能である。さらに、これらの XML 形式をそのまま提供することで、用語体系データの交換や再利用が可能となる。

#### 4.4 開発環境

本システムは、CGI の枠組みを利用して Web 上でユーザの要求に応じて動的に用語体系を可視化、閲覧できるシステムとして構築した。開発に利用した環境は以下の通りである。WWW サーバには Apache<sup>[11]</sup>を用い、CGI プログラムは Perl 5<sup>[12]</sup>で作成した。用語体系の検索などを高速に行うために Berkeley DB ライブラリ<sup>[13]</sup>および Suffix Array ライブラリ Sary<sup>[14]</sup>を利用したインデックスを作成した。グラフ構造のレイアウトおよび画像出力は、グラフ描画ソフトウェア Graphviz<sup>[10]</sup>を用いた。また、XML から各種表示形式への変換を行う XSLT 処理系には Libxslt<sup>[15]</sup>を利用した。

```
<?xml version="1.0"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
                version="1.0">
  <xsl:output method="text"/>
  <xsl:template match="/">
    digraph G {
      <xsl:apply-templates/>
    }
  </xsl:template>
  <xsl:template match="word">
    "<xsl:value-of select="@id"/>"
    [label="<xsl:value-of select="name"/>",style="filled",color="gray"];
    <xsl:for-each select="child">
      "<xsl:value-of select="@idref"/>" [label="<xsl:value-of select="."/>"];
      "<xsl:value-of select="../@id"/>" -&gt; "<xsl:value-of select="@idref"/>";
    </xsl:for-each>
    <xsl:for-each select="parent">
      "<xsl:value-of select="@idref"/>" [label="<xsl:value-of select="."/>"];
      "<xsl:value-of select="@idref"/>" -&gt; "<xsl:value-of select="../@id"/>";
    </xsl:for-each>
  </xsl:template>
</xsl:stylesheet>
```

図 9 dot 形式に変換する XSLT スタイルシート例

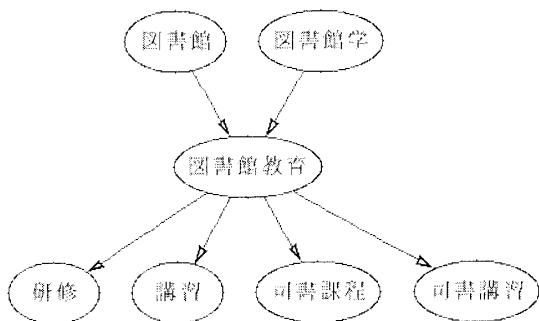


図 10 動的に生成した画像の例

#### 4.5 実行例

本節では、各機能を例とともに説明する。

HTML による一覧表示機能は、各見出し語の上位語と下位語を表示する機能である。表示された上位語と下位語はハイパーリンクになっており、ブラウザ上で語をクリックすることで、指定した見出し語の表示を次々で行うことができる。図 11 は NDC における用語「図書館」の持つ関係を表示しており、この中から「図書館活動」を選んでいところである。クリックすると図 12 のように「図書館活

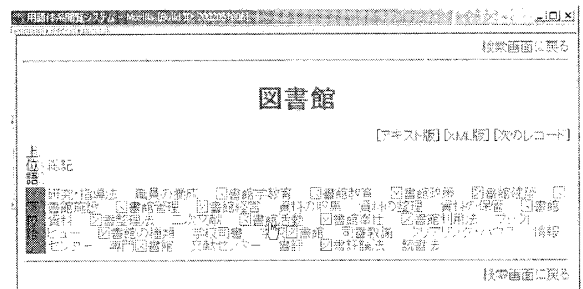


図 11 HTML による一覧表示: 「図書館」(NDC)

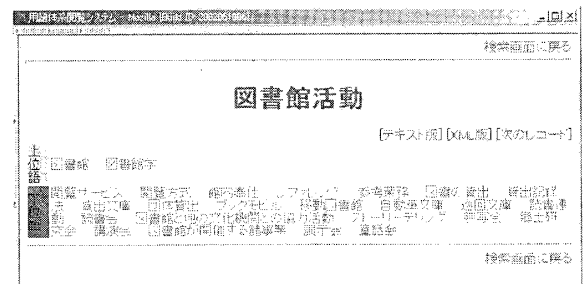


図 12 HTML による一覧表示: 「図書館活動」(NDC)

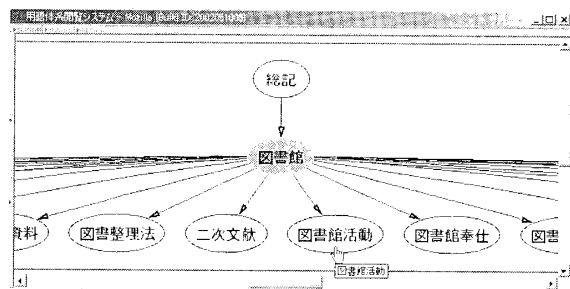


図 13 グラフ構造のクリックابلマップ表示: 「図書館」(NDC)

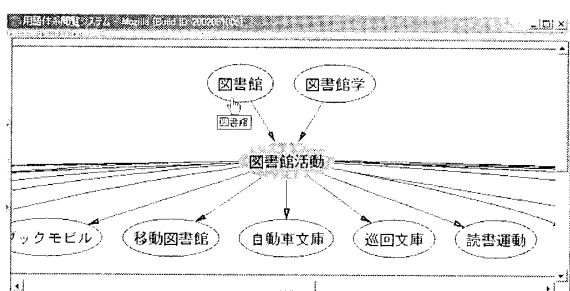


図 14 グラフ構造のクリックابلマップ表示: 「図書館活動」(NDC)

動」の上位語と下位語を見ることができる。

各用語体系のグラフ構造のクリックابلマップ表示機能は、各用語体系における見出し語をノードとし、上位階層の語を上、下位階層の語を下に配置したグラフ表現を表示する機能である。表示された上位語と下位語のノードはハイパーリンクになっており、クリックすることで、次々と指定した見出し語の表示を行うことができる。図 13 は「図書館」を表示しており、この中から「図書館活動」を選んでいくところである。このノードをクリックすると図 14 のように「図書館活動」についての上位語と下位語を見ることができる。

また、指定した語から上下に何階層まで表示するかも指定できる。図 15 は、ODP における見出し語「コンピュータ」からたどることができる 3 階層分の見出し語を同時に表示した例である。この複数階層の同時表示機能により、表示する範囲を広げて眺めることが可能になり、全体像の把握がより容易となる。

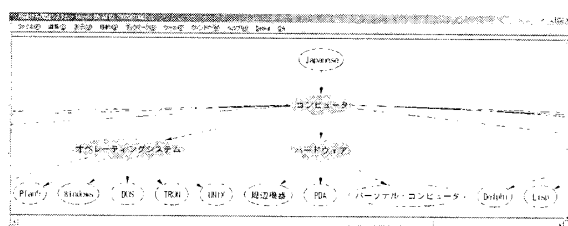


図 15 グラフ構造のクリックابلマップ表示 (複数階層同時表示): 「コンピュータ」(ODP)

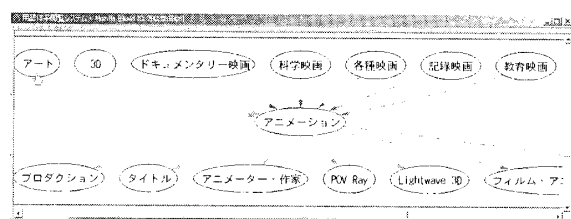


図 16 複数用語体系の重ね合わせ結果: 「アニメーション」

複数用語体系の重ね合わせ結果の表示では、異なる用語体系の関係が一目でわかるように、EDR は青、NDC は赤、ODP はオレンジ色というように、各用語体系での関係を表すエッジはそれぞれ別の色として表示している。また、一体系中で別の ID を振られている見出し語も、重ね合わされて表示される。図 16 は、複数の用語体系の重ね合わせ表示の例であり、図 8 と同じく EDR、NDC、ODP の 3 つの用語体系に存在する見出し語「アニメーション」の重ね合わせ結果に対して、グラフ構造のクリックابلマップ表示を行っている。ここで、見出し語「アニメーション」は、

- EDR における「フィルムアニメーション」の上位語
- NDC における各種映画の下位語
- ODP における「アート」の下位語
- ODP における「3D グラフィックス・ソフトウェア」の下位語

という、複数の体系における階層構造上の見出し語を重ね合わせたものである。このように重ね合わせを行うことで、その見出し語が持つ関係をより広い観点から一望できる。グ

**「エネルギー」にヒットした見出し語リスト**

**完全一致:**  
7件のヒットがありました。

| 用語体系 ID   | 見出し語  | 統合表示 | その他の表示形式                        |
|-----------|-------|------|---------------------------------|
| EDR 10838 | エネルギー | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| NDC 17011 | エネルギー | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| ODP216534 | エネルギー | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| EICnet 19 | エネルギー | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| NDC 13010 | エネルギー | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| NDC 782   | エネルギー | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| NDC 11321 | エネルギー | *    | [Text] [XML] [HTML] [GIF] [Doc] |

**部分一致:**  
127件のヒットがありました。

| 用語体系 ID   | 見出し語           | 統合表示 | その他の表示形式                        |
|-----------|----------------|------|---------------------------------|
| EDR 2712  | アークエネルギー       | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| EDR 3935  | アークエネルギー       | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| EDR 10839 | エネルギー資料情報相互連携性 | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| EDR 10841 | エネルギーアライアンス    | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| EDR 10843 | エネルギーアライアンス    | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| EDR 10854 | エネルギーギョウ       | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| EDR 10884 | エネルギーデータベース    | *    | [Text] [XML] [HTML] [GIF] [Doc] |
| EDR 10885 | エネルギーデータベース    | *    | [Text] [XML] [HTML] [GIF] [Doc] |

図 17 見出し語検索機能 (検索結果): 「エネルギー」

グラフ構造を表示する際には、その見出し語同士を結ぶエッジの色を変えることで、それぞれがどの用語体系に由来する関係であるかが直観的に判別できる。

また、見出し語検索機能では、各用語体系の見出し語を検索できる。見出し語の検索は、検索したい用語体系を複数選択することで、複数の用語体系を同時に横断的に検索できる。この見出し語検索では、入力された単語と完全一致する見出し語を表示するだけでなく、部分一致した見出し語も表示する。図 17 は「エネルギー」に対する検索結果である。検索結果は、完全一致した語、部分一致した語の順で表示し、ヒットした見出し語数と各見出し語毎に、その語の属する用語体系、見出し語 ID、見出し語、見出し語の階層関係の表示へのハイパーリンクを表示する。ここでは、EDR、NDC、EIC、ODP の 4 つの用語体系に一致する見出し語があることがわかる。また、部分一致した検索語については、一致した部分をわかりやすいように赤で強調表示している。この中から見出し語を選択してクリックすると、グラフ構造のクリッカブルマップ表示機能へ飛ぶことができる。

## 5 考察

本研究では、用語体系の作成、維持を支援するブラウジングシステムを構築した。本節では、本研究における用語体系の重ね合わせとブラウジングという 2 つの観点から、既存の関連研究との相違点と用語体系の拡張について述べる。

### 5.1 関連研究

用語体系閲覧システムの既存の研究には、後藤ら [16] や市丸 [17] によるシソーラスや用語辞書の概念体系をグラフィカルに表示し、ブラウジングするシステムがある。RDFViz [18] は、RDF データのグラフ構造を表示するシステムで、Web 上から利用できる。また、情報可視化に関する研究では、Pirolli ら [19] による階層構造を Focus+Context 手法を用いて可視化するシステムなどが開発されている。これらのシステムはいずれも、個別の体系を表示して閲覧する機能を持つが、本研究のように、複数の体系を重ね合わせて表示する機能はない。

また、複数のオントロジーを自動的に統合する試み [20][21][22] や複数言語の体系を統合する試み [23][24] も行われている。しかし、これらのアプローチでは、統一された一つの体系を構築することに主眼が置かれていて、複数体系を重ね合わせてその結果をブラウジングしながら取捨選択したり、再利用したりするという本研究のような観点からの言及はされていない。

一方、Noy ら [25] は利用者との対話によりオントロジーのマージを半自動化する手法を提案している。この手法に基づくシステムは、オントロジー構築用エディタ Protégé [26] 上で、PROMPT プラグイン [27] として実現されている。しかし、PROMPT は 2 つのオントロジーのマージを扱うもので、本研究のように同時に 3 つ以上の体系を重ね合わせることはできない。

## 5.2 用語体系の拡張

本研究で構築したブラウジングシステムの枠組みを使えば、複数の重ね合わせ手法を切り替えることも原理的には可能であると考えられる。例えば、用語の照合に依らず、異なる言語の体系同士を重ね合わせできる手法などが考えられる。Sowa は文献 [28]において、このような手法をアラインメント (alignment) と呼び、「オントロジー間での概念および関係のマッピング」と定義としている。本システムにおいて様々な体系同士を重ね合わせる際に、その重ね合わせ手法を拡張したり、このアラインメントの考え方にもとづく手法を切り替えて使うことで、より一層、複数体系間の相互利用が強まるものと考ええる。

また、ブラウジングシステムが動的に生成する XML の形式は、表示形式には依存していないため、新たな表示形式に容易に対応可能である。例えば、VRML (Virtual Reality Modeling Language) による 3 次元表示やグラフ表現のカスタマイズなどができると考える。

さらに、この XML 形式をそのまま提供することで、用語体系データの交換や再利用が可能となる。例えば、一部の用語の構造をそのまま取り出して自分の体系に組み込むなどが可能になるものと考ええる。

## 6 おわりに

本研究では、用語体系の再利用による作成支援を目指して複数用語体系のブラウジングシステムを構築した。分野や規模の異なる 5 つの用語体系を対象として、本システムがこれらに適用可能なことを示した。また、用語間の関係を分かりやすく表示し、その体系の構造を直観的に把握できるようになった。また、複数用語体系を重ね合わせて表示することで、用語の持つ様々な関係をとらえることができるようになり、それらの関係を新たな体系の構築に再利用することができる。

このシステムをさらに発展させることにより、様々な用語体系間での再利用が可能とな

り、様々な分野の用語体系構築に寄与するものと考えている。

## 参考文献

- [1] “ODP: Open Directory Project”, <http://dmoz.org/> (参照 2002-12-04).
- [2] 「EIC ネット」, <http://www.eic.or.jp/> (参照 2002-12-04).
- [3] Berners-Lee, Tim; Hendler, James; Lassila, Ora: “The Semantic Web”, *Scientific American*. May 2001, pp. 34-43, 2001.
- [4] 特集: 「セマンティック Web」, *情報処理*, vol. 43, no. 7, pp. 707-750, 2002.
- [5] Bray, Tim; Paoli, Jean; Sperberg-McQueen, C. M.; Maler, Eve: “Extensible Markup Language (XML) 1.0 (Second Edition)”, World Wide Web Consortium, 2000, REC-xml-20001006. <http://www.w3.org/TR/REC-xml> (最終更新 2000-10-06).
- [6] 伊東千夏; 高久雅生; 江草由佳; 大懸晶子; 宇陀則彦; 石塚英弘: 「用語体系の統合及び可視化システムの試作: 用語体系の補完及び相違点の発見を目的として」, *情報処理学会 情報学基礎研究会研究報告*, vol. 101, no. 65, pp. 1-8, 2001.
- [7] 日本電子化辞書研究所: 「EDR 電子化辞書 2.0 版仕様説明書」, 東京, 日本電子化辞書研究所, 1999, TR-006, <http://www.ijnet.or.jp/edr/>.
- [8] 日本図書館協会編: 「日本十進分類表新訂 8 版機械可読データファイル (NDC・MRDF8)」, 東京, 日本図書館協会, 1989, (電子媒体).
- [9] Clark, James: “XSL Transformations (XSLT) Version 1.0”, World Wide Web Consortium, 1999, REC-xslt-19991116, <http://www.w3.org/TR/xslt> (最終更新 1999-11-16).

- [10] AT&T: “Graphviz”, <http://www.research.att.com/sw/tools/graphviz/> (最終更新 2000-10-19).
- [11] Apache Software Foundation: “Apache HTTPD Project”, <http://httpd.apache.org/> (参照 2002-12-04).
- [12] Wall, Larry; Christiansen, Tom; Schwartz, Randal L (近藤 嘉雪訳): 「プログラミング Perl 改訂版」, 東京, オライリージャパン, 759p., 1997.
- [13] “The Sleepycat Software Homepage”, <http://www.sleepycat.com/> (参照 2002-12-04).
- [14] 高林哲: 「Sary: Suffix Array のライブラリとツール」, <http://sary.namazu.org/> (最終更新 2002-09-18).
- [15] Veillard, Daniel: “The XSLT C library for Gnome”, <http://xmlsoft.org/XSLT/> (参照 2002-12-04).
- [16] 後藤高信; 鈴木俊介; 後藤智範: 「階層構造を持つ用語データのための Browsing Tool」, 情報知識学会 第7回研究報告会論文集, pp. 57-60, 1999.
- [17] 市丸夏樹: 「シソーラスブラウザ xthes homepage」, <http://lang.is.kyushu-u.ac.jp/pub/xthes/>(参照 2002-12-04).
- [18] Brickley, Dan: “RDFViz”, <http://www.ilrt.bris.ac.uk/discovery/rdf-dev/rudolf/rdfviz/> (参照 2002-12-04).
- [19] Pirolli, Peter; Card, Stuart K.; Van Der Wege, Mija M.: “The Effect of Information Scent on Searching Information Visualizations of Large Tree Structures”, *Advanced Visual Interfaces 2000*, Palermo, pp. 161-172, [http://www.inxight.com/pdfs/white\\_papers/info\\_scent.pdf](http://www.inxight.com/pdfs/white_papers/info_scent.pdf) (参照 2002-12-18).
- [20] Ambite, José Luis; Arens, Yigal; Hovy, Eduard; Philpot, Andrew; Gravano, Luis; Hatzivassiloglou, Vasileios; Klavans, Judith: “Simplifying Data Access: The Energy Data Collection Project”, *Computer*, vol. 34, no. 2, pp. 47-54, 2001.
- [21] Hovy, Eduard; Philpot, Andrew; Ambite, José Luis; Arens, Yigal; Klavans, Judith et al: “Data Acquisition and Integration in the DGRC’s Energy Data Collection Project”, *Proceedings of the dg.o 2001 Conference*, 2001. <http://www.dgrc.org/dgrc/dgo2001/papers/session-2/hovy.pdf> (参照 2002-12-04).
- [22] Ichise, Ryutaro; Takeda, Hideaki; Honiden, Shinichi: “Rule Induction for Concept Hierarchy Alignment”, In *Proceedings of the IJCAI-01 Workshop on Ontology Learning (OL-2001)*, pp. 26-29, 2001. <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-38/ichise-IJCAI-OL.pdf>(参照 2002-12-18).
- [23] 荻野孝野: 「異なる言語の繋ぎ手としての多言語シソーラスへの試み」, *人工知能学会誌*. vol. 15, no. 4, pp. 567-574, 2000.
- [24] Asanoma, Naoki: “Alignment of Ontologies: WordNet and Goi-Taikei”, *Proceedings of the NAACL 2001 Workshop on WordNet and Other Lexical Resources*, Pittsburgh, 2001. <http://www.seas.smu.edu/rada/mwnw/papers/WNW-NAACL-233.pdf.gz> (参照 2002-12-18).
- [25] Noy, N. F.; Musen, M. A.: “PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment”. In the *Proceedings of the 17th National Conference on Artificial Intelligence*, 2000. SMI-2000-0831. [http://smi-web.stanford.edu/pubs/SMI\\_Abstracts/SMI-2000-0831.html](http://smi-web.stanford.edu/pubs/SMI_Abstracts/SMI-2000-0831.html).
- [26] Noy, N. F.; Sintek, M.; Decker, S.; Crubezy, M.; Fergerson, R. W.; Musen, M. A.: “Creating Semantic Web Contents with Protégé-2000”, *IEEE Intel-*

ligent Systems, vol. 16, no. 2, pp. 60-71,  
2001.

[http://www.jfsowa.com/ontology/  
gloss.htm](http://www.jfsowa.com/ontology/gloss.htm) (最終更新 2001-08-25).

- [27] Noy, Natasha: “The PROMPT Tab”,  
[http://protege.stanford.edu/plugins/  
prompt/prompt.html](http://protege.stanford.edu/plugins/prompt/prompt.html)(参照 2002-12-04).

(2002年12月26日受付)

(2003年2月26日採録)

- [28] Sowa, John F.: “Glossary”,

## 文化財情報の記述項目 — 現状の分析と系統化の課題 —

# Categories for Describing Information of Cultural Properties — Present State and Problems to be Solved for Systemizing —

田良島 哲\*

Satoshi TARASHIMA

伝統的な文化財の情報記述には独自の方法が蓄積されているが、組織的な教育が不十分であり、学術的な議論も少ないため、情報の共有が困難な状況が見られる。小論ではまず博物館等が刊行する図版目録において採用されている項目を紹介し、それを博物館データの国際標準案と対照しながら、同一と認められる項目の整理を試みた。項目の呼称はさまざまで、どの項目が使われているかは、分野によってばらつきが大きい。整理後の項目数は比較的少数にとどまり、基本的な文化財データについては博物館間で横断的に共有できる可能性は大きい。各項目は、依拠する学問によって記述方法に特色があるので、現在使われている主要な記述方法を紹介し、今後の情報共有を図る際に想定される問題点を指摘する。

Though unique styles have been accumulated for describing various aspects of Japanese traditional cultural properties, few education systems and academic discussions have been developed for it. Therefore much difficulty is caused in sharing information about cultural properties among museums and researchers. This paper introduces data categories used in published catalogs of some museums in Japan, and tries to arrange them into international museum data standard. As a result, despite names of the categories vary widely among museums, many share common concepts by appropriate arrangement. This fact suggests the possibility of sharing of basic object data among museums. As every category has its characteristic style of description based on academic background, the author introduces some typical styles and points out some problems expected in sharing information.

キーワード：文化財，目録，博物館，データ標準

cultural properties, cataloging, museum, data standard

## 1 問題の所在

美術工芸品・古文書・建造物など伝統的な文化財（以下「文化財」という）に関する情報記述は、それぞれの分野において独自の方法を持っている。しかし、これまで、その記述方法の伝承は大学や博物館において手工業的に行われてきたのが実状であり、図書館情報学のように、情報記述自体を対象とした系

統的な研究や教育はほとんど行われていない。古典籍や古文書は、書誌学や古文書学の蓄積があり、意識的ではないにせよ、情報記述に関する問題点が研究の素材としてとりあげられることがある。また、学会や教科書・参考書もその名をあげることができる。その一方で、美術史は、個別の大学教育でこそ「美術史調査法」「美術資料調査法」といった講座を見ることができるが、公刊された教科書は未だになく、このような問題について研究者間で議論が行われることも、ほとんどないように思われる。この点は、美術史学の側から、島

\* 文化庁文化財部

Department of Cultural Properties, Agency for Cultural Affairs  
stara@bunka.go.jp



尾新氏が

「日本美術史の教科書と言えるものは、研究者向けの『絵画調査法』などというものを含めて未だにほとんどなく、教育システムも確立しているとは言い難い。『科学的では割り切れぬものがある』ことは確かだが、『科学的で割り切れる』部分まで拒否する傾向も未だに残っている」

と指摘している点<sup>[1]</sup>からも裏付けられよう。文化財情報の記述については、次のような問題が予想される。

- \* 形態や伝来が多様であり、一律の記述項目設定が困難である。
- \* 記述が歴史的な条件に制約されるため、合理性だけで項目やスタイルを設定しにくい。
- \* 関係する学問分野が多岐にわたり、相互の検討が難しい。  
筆者は別稿<sup>[2]</sup>において、文化財の情報記述の問題を調書記述に関連して取り上げ、
- \* 調書に本来設定されるべき必須事項が欠如している場合がある。
- \* 用語や数値などの標準化の度合いがまちまちである。
- \* 美術史学特有のオリジナル中心主義による調査視角の偏向が見られる。

といった点を指摘したが、その解決の方向については若干の抽象的な提案をするにとどまった。

このような問題点がある一方で、文化財を取り扱う博物館・美術館等（以下「博物館等」という）の施設は1970年代以降著しく増加しており、そこから公刊される展覧会図録・調査報告書・研究論文も膨大な数にのぼっている。この蓄積を有効に活用し、文化財に関する情報を広く社会に還元してゆくためには、大学・研究機関・博物館等において共有を図る必要がある。

文化財情報の記述の標準化については、近年、いくつかの分野で、国際標準が提案されている。博物館資料関係では、国際博物館会議国際ドキュメンテーション委員

会(ICOM/CIDOC)のInternational Guidelines for Museum Object Information: The CIDOC Information Categories(IGMOI)<sup>[3]</sup>や策定中のCIDOC Conceptual Reference Model(CIDOC/CRM)、文書館関係では、国際文書館評議会(ICA)が策定した文書・記録の分野におけるGeneral International Standard Archival Description,ISAD(G)と、その電子化に関する規格案であるEncoded Archival Description(EAD)などがあげられる。今後、日本における情報記述も、これら国際標準案との整合性を意識しながら、検討してゆくことが求められよう。

小論は、これらの問題の解決について具体化を図るために、文化財目録で実際に用いられている記述項目を紹介し、多少の整理と問題点の指摘を試みたものである。まず、文化財情報として、博物館等における調査研究や文化財行政の現場でとりあげられる主要な記述項目について簡単な整理を行う。次にそれぞれの項目の実際的な記述手法に関する説明を加え、同時に現状での問題点について指摘する。専門家には自明と思われる点にも説明を加えているが、議論の素材として広い分野の関係者に理解を求めるためであり、諒とされたい。

対象は江戸時代以前のいわゆる美術工芸品とし、建造物や近代美術は含んでいない。なお、記述項目は基本的に単体の文化財に対するものであり、一括の古文書のような階層的な構成を持つ群としての文化財の記述は、必要とされる項目が異なってくるため、取り上げていない。

また、以下の記述で例示される文化財の具体的な名称は、主に国宝・重要文化財の名称を利用した。

## 2 文化財目録で用いられる記述項目

### 2.1 図版目録にみる記述項目

まず、文化財についての基本的な記述項目を実際の目録から抽出してみよう。数多くあ

るこの種の目録を網羅的に検討するのは、やや煩雑にわたるので、ここでは例示的に、文化財を多く所蔵する博物館・美術館の図版目録を取り上げた。図版目録は、文化財の所蔵館が所蔵品の写真図版と記述項目を掲載して公刊したものである。同種の目録類の中で最も公開性が高く、そこで取り上げられている項目は学芸員や研究者等の間で共有されるべき基本的情報と認識されていることから、検討の対象として適当と考える。

検討したのは、国立博物館（東京・京都・奈良）3館の分野別の図版目録14点、公私立で文化財を多く所蔵する館が刊行した図版目録6点の計20点である。結果は表1に示した。原則として、目録の「凡例」に記載された項目の呼称を相互に比較し、同種の概念を対応させたものである。これに国指定文化財の記述項目を定めた「国宝・重要文化財台帳規則」の項目、及び今後の国際的なデータ標準案との間のマッピングを考慮して、IGMOIが示すカテゴリとの対応を私案として示した。

## 2.2 結果

結果を見ると、すべての目録に共通した項目は、

名称（作品名・形態名）

法量（寸法）

の2項目のみである。他の項目は個々の目録によって、採用されている場合といない場合がある。

大多数の目録で採用されている項目としては、

品質形状（材質・形質・技法・構造等）

員数（数量）

作者（筆署名）

時代（世紀・製作年代等）

銘文（款記・賛者）

寄贈者（旧蔵・受理次第）

があげられる。作者では生没年を付記している目録がある。

また、関連情報として、

伝来（生産地・出土地）

番号（列品番号・館蔵品番号・受入番号）  
といった項目が含まれるものもある。

同じ国立館の目録間でも、記述項目がまちまちであるが、これは分野によって必須と認識される項目が異なることを示している。言い換えると、個々の目録作成者の関心が館内での情報の共有よりも、担当する専門分野内での利用に向いている、ということに他ならない。レジストレーションやドキュメンテーションの専門部門が弱体な日本の博物館等の組織形態から、十分予想される結果であろう。この点から、伝統的な文化財目録については、「形態や伝来が多様であり、一律の記述項目設定が困難」という悲観的な予想が裏付けられるようにも思われる。

しかしその一方で、表1の項目数は極端に多くなることはなく、「備考」というような漠然とした項目を含めて14項目にとどまっている。またIGMOIには22の基本カテゴリがあるが、それらとの間でもおおむね対応関係を確認することができる。

したがって、現在通用している図版目録の検討から帰納した結果を見ると、文化財の記述項目は、呼称がさまざま、どのような項目を採用するかについて、館や分野ごとにばらつきはあるが、概念を整理すると、一物件に対する記述項目の数は限定的である。この事実はまず、目録を作成する学芸員が、これくらいの項目数が基本的な公開情報として一般的だと認識していることを示している。同時に、適切なメタデータの設定によって、複数の博物館等を横断する項目の標準化を図ることが可能であることを示唆するものである。

## 3 記述項目に関する若干の分析

以下、前章で取り上げられた各記述項目について、典型的な記述の手法を説明し、併せて、それぞれの記述手法が持つ問題点について、特に情報記述の系統化・標準化を進めるという課題に留意しながら、指摘したい。電子的な処理に当たって想定される問題について

表1 博物館等の図版目録に採用された項目の比較

| IGMIOI:<br>The Information<br>Groups and Categories | 国宝重要文化<br>財台帳規則        | 東京国立博物<br>館図版目録<br>やまと絵 | 東京国立博物<br>館図版目録<br>日本彫刻 | 東京国立博物<br>館図版目録<br>仮面 | 東京国立博物<br>館図版目録<br>漆工調度<br>(文房具) | 東京国立博物<br>館図版目録<br>小袖服飾 | 東京国立博物<br>館図版目録<br>日本陶磁 | 東京国立博物<br>館図版目録<br>僧侶筆跡・唐<br>様 | 京都国立博物<br>館図版目録<br>絵画(中国・朝<br>鮮・日本桃山時<br>代以前) | 京都国立博物<br>館<br>藏品図版目録<br>彫刻・建築 |
|-----------------------------------------------------|------------------------|-------------------------|-------------------------|-----------------------|----------------------------------|-------------------------|-------------------------|--------------------------------|-----------------------------------------------|--------------------------------|
| Object Title Information / Object Name Information  | 名称                     | 1993年<br>作品名            | 1999年<br>名称             | 1970年<br>名称           | 1985年<br>名称                      | 1983年<br>形態名            | 1966年<br>名称             | 1995年<br>名称                    | 1989年<br>作品名                                  | 1993年<br>作品名                   |
| Part and Component Information                      | 員数                     | 員数                      | 員数                      |                       |                                  |                         |                         | 員数                             | 員数                                            | 員数                             |
| Material and Technique Information                  | 品質及び形状                 | 材質                      | 品質構造                    | 材質                    |                                  | 地質・模様・技<br>法            | 窯                       | (品質形状)                         | 材質                                            | 品質                             |
| Measurement Information                             | 寸法又は重量                 | 寸法(縦×横<br>cm)           | 法量(cm)                  | 法量(縦×幅<br>cm)         | 寸法(縦×横<br>×高さcm)                 | 法量(丈×桁<br>cm)           | 寸法(高さ・口<br>径・底径cm)      | 法量(縦×横<br>cm)                  | 法量(cm)                                        | 法量(cm)                         |
| Object Production Information (group/person name)   | 作者                     | 作者                      |                         |                       | 作者                               |                         |                         |                                | 筆者名                                           | 作者伝来等                          |
| Object Production Information (date)                | 制作の年代又は時代              | 時代・世紀                   | 製作時代・年<br>代又は世紀         |                       | 時代                               | 製作年代(時<br>代・世紀)         | 時代                      | (時代)                           | 制作時代                                          | 時代・世紀                          |
| Mark and Inscription Information                    | 商賈、奥書、銘<br>文等          | 歌記・賛者                   | 銘記                      | 銘文                    |                                  |                         |                         |                                |                                               |                                |
| Acquisition Information                             |                        | 寄贈者                     | 受理次第                    | 旧蔵・寄贈者                | 寄贈者                              |                         | 寄贈者                     | (寄贈者)                          | 寄贈者名                                          |                                |
| Object Number Information                           |                        |                         | 列品番号                    |                       | 列品番号                             | 列品番号                    |                         | 列品番号                           | 館登録番号                                         | 館登録番号                          |
| Object Association Information                      | 伝来その他参<br>考となるべき<br>事項 | 伝来その他                   |                         |                       |                                  |                         | 出土地                     |                                |                                               | 製作地                            |
|                                                     |                        |                         | 付属品                     |                       |                                  |                         |                         |                                |                                               |                                |
|                                                     |                        |                         | 納入品                     |                       |                                  |                         |                         |                                |                                               |                                |
|                                                     |                        |                         | 備考                      |                       |                                  |                         |                         |                                |                                               |                                |
| Reference Information                               |                        |                         | 文献                      |                       |                                  |                         |                         |                                |                                               |                                |

表 1 (続き)

| IGMOI:<br>The Information<br>Groups and Categories           | 京都国立博<br>物館蔵品図<br>版目録<br>陶磁・金工 | 京都国立博<br>物館図版目<br>録<br>日本書跡 | 奈良国立博<br>物館蔵品図<br>版目録<br>絵画 | 奈良国立博<br>物館蔵品図<br>版目録<br>彫刻 | 奈良国立博<br>物館蔵品図<br>版目録<br>工芸(仏教<br>工芸) | 大阪国立博<br>物館蔵品選<br>集  | 福岡市美術<br>館所蔵品目<br>録<br>古美術 | 石川県立美<br>術館所蔵品<br>図録 | 根津美術館<br>列品図録 | 静嘉堂宝鑑       | 大和文華館<br>所蔵品図版<br>目録<br>絵画書跡<br>(日本編) |
|--------------------------------------------------------------|--------------------------------|-----------------------------|-----------------------------|-----------------------------|---------------------------------------|----------------------|----------------------------|----------------------|---------------|-------------|---------------------------------------|
| Object Title Informa-<br>tion / Object Name In-<br>formation | 1989年<br>作品名                   | 1983年<br>作品名                | 1988年<br>名称                 | 1989年<br>名称                 | 1992年<br>名称                           | 1986年<br>作品名         | 1992年<br>作品名               | 1994年<br>作品名         | 1984年<br>名称   | 1992年<br>名称 | 1990年<br>作品名                          |
| Part and Component<br>Information                            | 員数                             | 員数                          | 員数                          | 員数                          |                                       | 員数                   | 員数                         | 員数                   | 数量            | 員数          | 員数                                    |
| Material and Tech-<br>nique Information                      |                                | 形質                          | 品質形状                        | 形状                          | 品質技法                                  | 材質                   | 品質形状                       |                      | 形質            | 品質形状        | 材質                                    |
| Measurement<br>Information                                   | 法量 (cm,<br>g)                  | 法量 (縦×<br>横 cm)             | 法量 (cm)                     | 法量 (cm)                     | 法量 (cm)                               | 寸法                   | 法量                         | 寸法                   | 法量 (cm)       | 法量          | 法量                                    |
| Object Production<br>Information<br>(group/person name)      | 作者                             | 筆者名                         |                             | 作者                          |                                       | 作者及生没<br>年           | 作者 (生没<br>年)               | 作者 (生没<br>年)         | 作者            | 作者          | 筆者 (生没<br>年)                          |
| Object Production<br>Information (date)                      |                                | 時代                          | 時代                          | 時代                          | 時代                                    | 製作年代                 | 時代・制作<br>年                 | 制作年代                 | 時代            | 制作時期        | 時代・制作<br>年                            |
| Mark and Inscription<br>Information                          | 銘文                             |                             | 銘文                          | 銘文                          | 銘文                                    |                      | 費者                         |                      |               |             |                                       |
| Acquisition<br>Information                                   | 寄贈者名                           | 寄贈者名                        |                             |                             |                                       | コレクショ<br>ン名・寄贈<br>者名 | コレクショ<br>ン・收藏方<br>法 (收藏年)  | 收藏経過                 |               |             |                                       |
| Object Number<br>Information                                 | 館登録番号                          | 館登録番号                       | 館蔵品番号                       | 館蔵品番号                       | 館蔵品番号                                 |                      |                            | 受入番号                 | 調書検索番<br>号    |             | 登録番号                                  |
| Object Association<br>Information                            | 生産地・出<br>土地                    |                             | 伝来                          | 伝来                          | 伝来等                                   |                      |                            | 伝来                   |               |             |                                       |
|                                                              |                                |                             |                             | 像内納入品                       |                                       |                      |                            |                      |               |             |                                       |
|                                                              |                                |                             |                             |                             |                                       |                      |                            |                      |               |             |                                       |
| Reference Information                                        |                                |                             |                             | 参考文献                        | 参考文献                                  |                      | 注記                         |                      | 略解            | 解説文         | 参考文献                                  |

でも若干とりあげた。

なお、各項目の統制語に当たる語は、私案であり、IGMOI のカテゴリに対応する適切な表現と考えたものである。

### 3.1 名称

#### *Object Name Information*

#### *Object Title Information*

「名称」と総括した項目には、絵画の画題や仏像の尊名、工芸品の種類などを含んでいる。日本の文化財の名称を検討してみると、IGMOI では二つのカテゴリにまたがるのがわかる。国立博物館の目録の中で「形態名」という項目になっているのが、object name の概念に近く、「作品名」が object title に該当する。具体的には工芸品の「梵鐘」は object name であるが、絵画の「四季花鳥図」「阿弥陀三尊来迎図」は object title ということになる。

分野によっては、他の項目に記入される内容が「名称」として取り扱われることが少なくない。書跡や古文書では、

無学祖元墨蹟  
織田信長書状

のように、「作者」に属する情報が名称に入り、単に「墨蹟」「書状」という名称は使われないのが一般的である。

また、工芸品の場合、

白地松鶴亀草花文繡箔肩裾小袖（染織）  
蓬萊山蒔絵硯箱（漆工）

のように文様や技法など品質・形状に属する事項を名称の中に織り込んでしまう命名法が普及している。この名称の例では

[文様] + [技法] + [品目]

白地松鶴亀草花文 + 繡箔 + 肩裾小袖

蓬萊山 + 蒔絵 + 硯箱

と分解される。表記のスタイルとしては、

肩裾小袖（白地松鶴亀草花文繡箔）

硯箱（蒔絵 蓬萊山文様）

でもさしつかえない。

仏像の「立像」「坐像」「半跏像」も概念的には形状に属するが、

十一面観音立像

弥勒菩薩半跏像

のように、名称に含めてしまうのが一般的である。

「品質・形状」の項目でも再説するが、このような複合的な名称を電子的に処理する際には、将来的にはソーラスなどに基づいて、名称の中から品質・形状に関する用語を切り出す工夫も必要になるかもしれない。

### 3.2 員数

#### *Part and Component Information*

文化財の世界では「名称」と「員数」を組み合わせて、使われる場合が多い。

伝統的な文化財では、仏像の「阿弥陀三尊」「十二神将」「二十八部衆」、卷子装の「法華経」ならば8巻あるいは10巻、「大般若経」ならば600巻（帖）というように、標準となる員数が決まっていることがあり、名称と員数を組み合わせることは、標準のセットにかなっているかどうかを確認する意味もある。

日本語の員数には、助数詞という、ある意味では便利な、ある意味ではめんどうなツールがある。適切な助数詞を用いれば、形態をある程度知ることができる。絵画や書跡で「1幅」は掛幅装、「1巻」は卷子装<sup>[4]</sup>、「1合」といえば蓋のある器物、といったぐあいである。しかし、適切な助数詞を決定できなければ、かえって誤解を招きかねず、使いこなすためには経験が必要である。器物と助数詞の関係は比較的例外が少ないので、ある程度の標準化がなされれば、たとえばデータベース上で器物を決めると、助数詞が自動的に対応する、というような処理が可能になるかもしれない。

員数では、項目に示されている数値が、実際の数を示さないことがある。具体的には、a pair ofを示す「一对」（例：仁王像・狛犬）「一双」（屏風）は「2 軀」「2 隻」としないのが慣例であり、データベースで点数の集計をする際などには設計上留意しなければならない。また甲冑・鞍・密教法具などに使われる「一具」という表記は、具体的な数ではなく本来

あるべきセットが揃っているという意味であり、実際の数には物件の種類による。

### 3.3 品質・形状

#### *Material and Technique Information*

「品質・形状」という語は、文化財に限らず、各種産業においてもよく使われる。たとえば、あるインターネット上の用語集の日本工業規格 (JIS) についての説明<sup>15)</sup>に、

JIS (中略) 工業標準化法にそって工業製品の品質、形状などの規準を定めたものをいう。

と要約しているのがその例である。人工的な器物の属性を記述する基本的概念の一つと言ってよいであろう。

各博物館の目録でもさまざまな用語が使われているように、文化財において品質・形状という概念の定義はかなりむずかしい。a. 素材 b. 技法 c. 様式 d. 形態 e. 文様などかなり幅の広い内容を含んでおり、特定の事項だけに絞ることはできない。既述のように、技法や文様などは名称に織り込んでしまう場合もあり、どこまでを品質、形状、どこからを名称に含めるかはある程度判断のわかれるところである。IGMOI の定義もかなりゆるやかであり、厳格な運用を想定しない方がよいのであろう。

通常は

紙本墨画 絹本着色 (絵画の場合)

木造 銅造 乾漆 (彫刻の場合)

というように素材、あるいは素材+技法を組み合わせた表記がなされる。

文化財は器物の種類によって素材や技法が特定される場合が多いので、目録記述の中で素材や技法が省略されることもある。たとえば、漆工のベースとなる素地は、ほとんどが木である。その場合、目録レベルでは素地に関する記述はなく、ただちに「蒔絵」「沈金」といった技法を記述するのが一般的である。素地が紙や皮革であれば、それぞれ「紙胎」「漆皮」という用語が特記される。

技法に選択の余地がない場合も、その技法についての目録記述が省略されることがある。梵鐘は、製作技法としては「鑄造」になるが、梵鐘製作の技法は鑄造のみであるから、これも目録レベルではわざわざ特記しないことが多い。ややくわしい説明がある場合は「鑄銅製」といった技法に関する記述が行われる。

絵画や書跡・典籍のように名称が主として文化財の内容を記述しているものは、

六曲屏風 (絵画)

折本装 (典籍など)

のように別に形態の記述をしておかないと、情報として不十分である。

品質・形状は原則的には客観化できるデータであり、解釈の余地なく一意に決まることがのぞましい。

### 3.4 法量

#### *Measurement Information*

「法量」とは聞き慣れない言葉かもしれないが、美術品の展覧会図録や文化財調査報告書にはよく使われる。日常的にいうところの「寸法」である。本来は仏像の寸法を表すもので、『日本国語大辞典』では、

仏像の丈量。立高と坐高とがあり、髪際から測るのを慣習とし、その高さによって丈六・半丈六・等身などの称がある。

と、説明している。しかし、第二次大戦前の宗教団体法施行令に

三 宝物ニ在リテハ種類、名称、員数、品質、形状、法量、作者及由来 (第7条)

などとあるように、以前から、仏像以外の宝物 (文化財) にも使われる場合が多く、逆に、インターネットの検索エンジンによる検索結果を見ると、基本的に文化財以外に関して使われている例は見られないようである。

文化財の含む情報の中で、法量は欠くことができない。せいぜい人間が手で持てる範囲の大きさである図書と異なって、大きさが多様だからである。たとえば仏像をとっても、大

は高さ 15 m 近い奈良の大仏から、小は 10 cm にも満たない金銅仏までである。法量のない文化財の記述は、全く用をなさないの、文化財の形姿を描写する上での基本的な情報の一つと言ってよい。

法量を記述する際は、a. 対象となる文化財のどの部分を測定しているのか、b. 単位が何かという二点が必須の項目である。この点は IGMOI のカテゴリが、

#### *Dimension Measurement*

##### *Measurement unit Measured part*

の 4 つのサブカテゴリに分けられていることが参考となろう。これに加えて、前稿で指摘したように<sup>[6]</sup>Dimension についてどれが優先的に記述されるべきかを確定しておく必要がある。

測定箇所は、絵画や書跡ならばオリジナルの部分である「本紙」（縦 × 横）が基本的な測定の対象であり、表具などは変更の可能性があるので、積極的に測定されない。彫刻も台座・光背・着装した宝冠などは一次的な測定対象としないのが通例である。同じように対象物件によって伝統的な測定箇所があるが、どの程度標準化がなされているかは、分野さらには個別の器物の形態によって様々であり、今後詳細な研究の必要がある。仏像の法量測定について、研究方法の変化との関連で津田徹英が論じているのが参考となろう<sup>[7]</sup>。

測定の単位は、現在では cm を単位とするのが一般的である。しかし、実地の調査においては仏像について言及したように丈・尺・寸を意識する必要がある。法量の測定は現物と目録との同定のために一定の精密さは必要であるが、同時に伝統的なモジュールとの対応を確認する意味合いが大きい。実際のところ保存状況の悪い文化財について、厳密な測定が困難であることは、常に意識しておかなければならない。

### 3.5 作者

#### *Object Production Information group/person name*

文化財は時代がさかのぼるほど作者が知られない場合が多く、また、分野や時代によっては「作者」という概念自体が無意味なこともある。

落款や銘文がある場合は、それにしたがって、

南蛮人渡来図 狩野内膳筆（絵画）

阿弥陀如来立像 快慶作（彫刻）

などと、明快に作者を書くことができる。また、集団的な製作物、たとえば梵鐘の場合、「大工物部季重」というように統括者の名前が銘文などで明確な場合は、これを作者として扱うことが可能である。

しかし、中世以前の陶磁器・染織品・漆工品等、多くの美術工芸品は無名であり、目録上「作者」を特定できない場合が多い。「作者」に準じる用語として、絵画における「狩野派」、刀剣における「一文字派」のような「流派」、陶磁器の目録に見える「窯」（瀬戸・信楽・丹波）などがあげられる。これらは object production information のサブカテゴリ group name や production place によく対応している。とりあえず「作者」として整理したこの項目について、将来的に情報共有を図る際には、より幅のある概念を採用した方が实际的であろう。

なお、名称にうたわれている名前と実際の制作者が異なるケースについても注意しなければならない。古文書で、たとえば、書状形式をとった天皇の命令書である「後醍醐天皇綸旨」と言えば、実質的な発給者は天皇だが、実際に書状を書いた差出人は「奉者」と呼ばれる側近の貴族である。この場合、古文書学では奉者を「差出人」として扱うのが通例で、文書の作者を後醍醐天皇とすると誤りになる。このような問題はそれぞれの学問分野に存在するので、十分な研究が必要である。

### 3.6 時代

#### *Object Production Information date*

時代は、奈良・平安・鎌倉…といった時代名で記述される場合と、これに客観的な年代を加える場合がある。銘文などで成立年代が

確定できる文化財については

鎌倉時代 弘安 4(1281) 年

といった表記をするのが通例であり、年代が推定になる場合は、

室町時代中期 15 世紀

というように記して、時代の幅を絞ることに  
なる。

一見客観化できそうな項目であるが、美術史では実年代と様式上の年代が一致せず、同じ名称で呼ばれる時代の幅が異なることがある。

日本史学では、桃山時代は関ヶ原の戦い(1600)頃までを指すが、美術史学では慶長末年(1615)ころまでを「桃山時代」としており、混乱を生じることがある。また、刀剣のように「上古」(歴史学では古墳時代に該当)といった独自の時代区分をしている分野もある。いずれもそれぞれの学問に一定根付いた用法であり、整理することは意外に困難である。

### 3.7 銘文等

*Mark and Inscription Information*

絵画ならば「画賛」「落款」、彫刻や工芸品ならば「銘文」、典籍ならば「奥書」「刊記」などと用語はさまざまであるが、当該文化財に付随して制作や伝来を示す文字記録を指す。画面上・像内・表面等に直接記載されているものを言う場合が多いが、仏像の納入品に記されている文章もこれに含める場合もある。

文化財の指定名称では、

- \* 嘉暦元年清拙正澄の賛がある
- \* 「道有」の印がある
- \* 像内に東大寺八幡宮安置、建仁元年十二月廿七日開眼、造立施主快慶の銘がある

といった要約した表記を取る。博物館等の目録ならば、

(台座天板裏面墨書銘)

「南都興福寺 大仏師法眼康俊 正中三年四月日」

というように、全文とそれが記載されている場所を記録しておくのが一般的であろうが、

インデックスとして簡略な表記を作っておくのも、検索等の際に有用である。

### 3.8 取得方法

*Acquisition Information*

各博物館の目録とも多くが「寄贈者」という呼称を用いているが、購入や出土もありうるので、正確にはいくつかの目録で使われている「受理次第」「取得経過」あるいはIGMOIの呼称にしたがって「取得方法」と言うべきであろう。

### 3.9 館蔵品番号

*Object Number Information*

博物館等が付与している番号は、公刊された目録でも掲載されていないことがままある。その場合は、目録作成者側に、番号は内部的な管理に利用できれば十分で、館外との情報交換には名称があれば間に合うという意識があるのかもしれない。しかし、今後の情報交換を考えると、特定の物件について一意に決まる項目として公開することがのぞましいと考える。

## 4 結論

博物館のドキュメンテーションについて議論をする機会があると、博物館等の目録に採用されている記述項目は、あまりに多様で整理や標準化は困難なのではないかという、漠然とした危惧が語られることがよくある。ほんとうに整理しきれないほど多様なのかという疑問から、小論のような簡単な試みを行って見た。

その結果によると、博物館等が公開情報として必要と認識し、現在刊行物で流通している程度の情報であれば、多様な呼称と内容を持つ項目の標準化は可能と思われる。小論では限られた数の博物館等のみを検討の対象としたが、伝統的な文化財に関しては代表的な施設を選んだので、一定の規範性はあるもの



と考えている。

今後、具体的には、国際的なデータ標準へのマッピングを足がかりとして、目録の項目の実例を多数比較検討し、項目の持つ概念の整理を行うと同時に、データ標準自体の適切さを検証してゆく作業が必要であろう。

小論では目録の形式に絞って問題点を紹介したが、次に導き出される実践的な課題として、記述に使われる用語の系統化・標準化があげられる。目録の形式面での共有は、情報共有の上で大きな前進となるが、そこで使われる用語とその組織化は、研究の進展や文化財の活用に一層寄与することが期待される。これは膨大な用語の詳細な分析が必要となり、論点も多岐にわたるため、稿を改めて論じることとしたい。

## 参考文献

- [1] 鳥尾新:『『雪舟研究史』雑感』、『日本における美術史学の成立と展開 (平成9~12年度科学研究費補助金基盤研究(A)(2)研究成果報告書, 研究代表者米倉迪夫)』, pp.410-419, 2001
- [2] 田良島哲:「古文化財の情報記述に関する問題点と展望」、『第2回アート・ドキュ

メンテーション研究フォーラム「美術情報の明日を考える」報告書』, pp.62-69, 2000

- [3] 全文は, <http://www.willpowerinfo.myby.co.uk/cidoc/guide/guide.htm> を参照。なお, 村田良二のサイト <http://www.ima.fa.geidai.ac.jp/~ryoji/museuminfo/index.html> には IGMOI の日本語訳とともに CIDOD/CRM の解説がある (2003年1月15日参照)。
- [4] 『日本目録規則 1987年版 改訂版』和図書適用細則は「1軸」を推奨しており, 文化財の立場からはやや疑問が残る。
- [5] 「ZDNet 情報処理技術者試験合格ヘルパーオンライン用語集」  
<http://www.zdnet.co.jp/magazine/techinfo/dictionary/1090.html> (2003年1月15日参照)
- [6] 田良島:注2論文
- [7] 津田徹英:「日本彫刻史研究における仏像の法量計測をめぐって」, 注1報告書, pp.384-409, 2001

(2003年1月17日受付)

(2003年2月4日採録)

# 「日本古典文学本文データベース」の利用調査

## An Investigation Report on the usage of the Full-text database for Japanese Classical Literature

山西 史子\*

Fumiko YAMANISHI

本稿は「日本古典文学本文データベース」の利用調査に基づき、下記の課題の評価を試みたものである。

- (1) 「日本古典文学本文データベース」の使われ方。
- (2) 「日本古典文学本文データベース」が利用者の研究活動に与えた影響。

調査方法は、電子メールを用いたアンケートによった。調査項目は10項目とし、その内容は、「日本古典文学本文データベース」の利用方法および、その研究活動への影響についてである。調査結果からは以下のような興味深い結果が得られた。

- (1) 「日本古典文学本文データベース」による調査を前提とした研究が行われている。
- (2) 研究のスピードを上げたり、生産性を高める影響は出ている。

また、利用者が、より多くの作品の収録や精確な本文の提供を求めていることも明らかになった。

This paper attempts to evaluate the following areas based on the usage survey performed regarding the Full-text Database for Japanese Classical Literature.

- (1) How the Full-text Database for Japanese Classical Literature is used
- (2) What effect the Full-text Database for Japanese Classical Literature has had on the research activities of its users

The survey was conducted using an E-mail questionnaire. Ten areas for survey were prepared, focusing on how the Full-text Database for Japanese Classical Literature is used and what effect it has had on research activities. We were able to obtain the following fascinating results from analysis of the survey results.

- (1) Research which presupposes use of the Full-text Database for Japanese Classical Literature is being conducted.
- (2) Research is being carried out faster, and productivity has also improved.

In addition, it has become clear from this survey that many users want to see a greater variety of works and more accurate texts available.

キーワード：日本古典文学本文データベース, 国文学研究資料館, 日本文学研究  
Full-text Database for Japanese Classical Literature, The National Institute of Japanese Literature  
Research of Japanese Literature

### 1 はじめに

人文科学の研究にコンピュータを導入する方法は1960年代から始まっている[1]。わが国においては、人文系研究現場でのパソコン

利用の普及は、1990年代に入ってからという指摘がある[2]。文献[2]では、近年の傾向について、歴史であれ文学であれパソコン抜きの研究はおおよそ想像しがたいと述べている。

このような背景を踏まえた上で、近年では人文科学の研究におけるコンピュータ利用の

\* 愛知淑徳大学大学院 文学研究科 図書館情報学専攻  
989045x@asu.aasa.ac.jp

実態やその研究への影響の調査もなされるようになってきている。歴史学の研究者に対する調査では、回答者の6割強がWWWを利用し、約4割が史料目録データベースやデジタル化された史料を入手しているという指摘もある [3].

国文学・国語学の研究への影響を考察する雑誌の特集や、文献も既に発表されている。「文学・語文」では、情報技術は文学研究をいかに変えるかという題で特集号を組んでいる。この中には、情報技術によって文学研究が変わるかどうかが明確な結論は見いだせないという意見も掲載されている [4]. 一方で、資料をデータとして扱う場合の処理や、インターネットによるコミュニケーションや情報の共有化によるコラボレーションの可能性の指摘もされていた [5]. 本稿ではこのような背景を踏まえた上で、「日本古典文学本文データベース」の利用者の調査を行うことで、資料の電子化の人文科学研究への影響を検討する。

## 2 調査の前提

### 2.1 項目設定の要件

「日本古典文学大系」を電子化した「日本古典文学本文データベース」の利用の実態を調査することにより、以下の2点について検討する。

- (1) 「日本古典文学本文データベース」がどのように利用されているか。
- (2) 「日本古典文学本文データベース」の利用により研究に影響がでてきたか。

(1) では、主に利用するデータベースの機能や具体的な利用目的を問う項目を設定し、調査を行う。(2) では、研究に使用する資料の変化や研究に影響が出てきたと感じているかを問う項目を設定し、調査を行う。

### 2.2 「日本古典文学本文データベース」の概要

「日本古典文学本文データベース」は、岩波書店が1957年～1969年にかけて刊行した「日本古典文学大系」全100巻の本文を電子化したものである。1つの作品を1つのファイルにすることを原則とし、約600作品を収録している。1999年からインターネット上で試験公開を開始しており、2002年2月時点で、利用登録者数は約1400名、1日平均約100件の利用がある。

古典本文のデータには国文学研究資料館が開発したマークアップ言語であるKOKIN規則でタグが付けられている [6]. KOKIN規則は、現在SGML/XMLへの変換が進められており、一部の作品については、XMLデータ、DTDデータを試験公開している [7] [8]. 知的財産権などの制約から、現在利用登録制となっており、以下の要件に該当する場合に利用が認められる [9].

- (1) 学校の教員及び調査研究機関の研究員並びに図書館職員
- (2) 大学の学生及び大学院の学生
- (3) 特に館長が適当と認めた者

利用者の研究テーマによる利用の制限は無く、研究目的であれば自由に利用できる。日本古典文学本文データベースの機能としては、作品(ファイル)を選択した後に、データの源である日本古典文学大系の書誌的データを確認したり、本文をダウンロードしたり、文字列検索を行う。

利用の多い作品(ファイル)は、「今昔物語集」「日本書紀」「源氏物語」「太平記」「宇津保物語」「栄花物語」「平家物語」「古事記」, 「万葉集」などである [10] [11].

## 3 調査

### 3.1 調査方法

調査方法は、電子メールを用いたアンケート

トによった。

調査対象：日本古典文学本文データベースの

利用登録者中，利用頻度の上位 312 名

調査手順：2002 年 2 月 20 日に質問内容を送

付し，3 月 10 日を回答期限とした。

日本古典文学本文データベースの利用登録者中，利用頻度の上位 312 名は，2000 年 4 月～2 月の間に 9 回以上データベースにアクセスした利用者である。利用者全体から，平均して 1 月に 1 度以上利用する高頻度利用者を抽出することにより，データベースを積極的に利用している研究者を調査対象とした。

### 3.2 調査内容

質問は日本古典文学本文データベースに関する 10 の設問で構成した。問 1～4 が，利用に関する質問，問 5～7 は研究動向への影響に関する質問，問 8 は日本古典文学本文データベースに今後期待すること，問 9 は同種のデータベースでよく利用するもの，問 10 がその他の意見とした(付録参照)。また，問 3，問 7 以外には「その他」や自由記述欄を設けている。

## 4 調査結果

### 4.1 回収結果と回収者の属性

回収結果は，有効回答数が 94 件 (30%) であった。

利用者には，利用登録時に，所属や研究テーマをたずねている。これらのデータから利用者の研究分野と所属・職業を識別すると以下のようになる。

所属・職業は，研究者・教育者が 59 名，学生(大学院生)が 29 名，その他が 9 名となっている。

利用の目的と研究分野は，国文学が 43 名，国語学が 26 名，歴史学が 13 名，中国文学と宗教が 2 名ずつ，その他・不明が 7 名，教育目的が 4 名であった。

ここからは，日本古典文学本文データベースが国文学・国語学を中心に様々な分野の研究と教育に使われていることが分かる。

### 4.2 日本古典文学本文データベースの利用目的と利用方法

回答者は日本古典文学本文データベースの利用規程により，研究者，学生，教員のいずれかに制限されている。また，目的も研究利用に限定されている。しかし，教員の利用者もいるため，利用目的が教育である可能性も考慮する必要があった。教育目的の回答者は 4 名あった。

利用方法に関しては，日本古典文学本文データベースの「資料としての位置づけ」「主に使用する機能」「使用する機能とその目的」の 3 つの間を設けている。

「主に使用する機能」では，「全文表示・ダウンロード」と「語彙・文字列検索」が全体の 9 割以上を占めていた。しかし「全文表示・ダウンロード」の機能を用いた場合，それを自分自身で加工する目的とそれ以外の目的の可能性も考慮しなければならない。同様に，「語彙・文字列検索」の機能に関しては，特定の場所を探しその部分に目を通す場合と，用例収集目的と二つの可能性がある。それらを識別するために「使用する機能とその目的」を調べる必要がある。

「使用する機能とその目的」の回答からは，回答者の 8 割近くが「用例収集のために語彙・文字列検索」を行っていることが分かった。同時に，「特定の場所を読むための語彙・文字列検索」と「本文データを加工するためのダウンロード」も回答者の半数以上が行っていた。

## 5 日本古典文学本文データベースの利用と研究への影響

日本古典文学本文データベースの研究上の有効性とその具体的成果の設問では，回答者の 8 割以上が「役に立った」と回答している。

研究成果の記入者は34名で、日本古典文学本文データベースを使用して作成した論文名の回答もあるが、より具体的に何を調べたのかの記入もあった。以下のような例である。

## 5.1 研究への具体的な利用

### 事例 A

「古代の公文書中で紫菜と表記されていたアマノリが平安時代に源順にとって甘苔と書かれ、次いで、うつほ物語の中で甘海苔と表記されて今日に至ったこと、同時代あるいはそれ以前に記録がないことが確認できた。」

### 事例 B

「例えば、日本で中世の自殺の研究をしているので、『切腹』だけではなく『捨身』『生害』『殺水』などの表現でこの現象が現れているので、『日本古典文学データベース』に載っているテキストでは関係のある表現で簡単に調べられることは、非常に役に立つ。」

### 事例 C

「『海鼠』という言葉は、日本古典文学において、どの程度、どこに出てくるのか、調べることができた。」

## 5.2 研究一般への影響

「研究への変化」の設問は、データベースの利用により研究全体への変化に対する主観的な感想を求めた問である。これに対しては、「変化なし」を回答した11名を除いて、何らかの変化を意識していると思われる。「コンピュータの利用増加」と「研究成果が出しやすくなった」の選択者がそれぞれ約4割を占めている。また、「今までとは異なったアプローチの研究ができるようになった」に対しても約3割の選択者がいた。具体的な記述としては、以下のようなものがある。

### 事例 D

「量的な観点から見ると、調査対象に劇的な増加を見、大きな流れをフォローしやすくなった。」

### 事例 E

「引用が大変楽になった。また索引を引くより、語句検索をかけた方が楽にまた正確に抽出でき、重宝している。」

### 事例 F

「手元に所持しない作品や、所持していてもコピー等の準備がしにくいテキストを、ダウンロードして加工すればすぐに教材として利用できるのは、大変便利であり、教育方法の改善の上で成果が上がっている。」

### 事例 G

「とくに近世のテキストは語彙索引が刊行されておらず、これまでは用例の検索に多大の時間を要したが、日本古典本文データベースのおかげで近世の用例にあたりやすくなった。」

事例 D と E からは、研究の効率への日本古典文学本文データベースの影響が読みとれる。事例 F からは、冊子体資料の代用品として、事例 G からは、出版物として刊行されていない資料の代用品としてデータベースが使用されていると思われる。

## 5.3 資料利用への影響

「研究における資料利用の変化」の設問は、日本古典文学本文データベースが研究資料として、冊子体資料とどのような関係にあるのか、冊子体資料の利用に影響を与えるのかを目的とした問である。これに対しては、「冊子体よりもデータベースを利用するようになる」が52名と最多であったが、「データベースよりも冊子体を利用するようになる」の回答者も28名あった。

## 5.4 日本古典文学本文データベース に対する要望

この問は利用者の要望を探るために設定している。「特になし」を選択した1名以外は、何らかの要望があることになる。「より多くの作品の収録」は89名が、「論文・図書とのリンク」は46名の回答者が選択していた。「外字が扱えるようにしてほしい」は45名、「機能の充実」は25名の回答者があった。13名が選択していた「その他」の回答と併せて、後述する。

## 6 考察

### 6.1 日本古典文学本文データベース の利用目的

回答者の属性と「データベースの利用目的」からは、日本古典文学本文データベースが、国文学、国語学、歴史学の研究者に研究目的で利用されていることが明らかである。また、「データベースの有効度」からは、日本古典文学本文データベースに好意的な利用者がアンケートに回答してくれたことも明らかである。

「利用機能」「利用目的」からは、電子化されたテキストデータの特質である検索機能を活用し、「用例収集」を行ったり読む場所を特定するために使用していることが分かる。また、「本文をダウンロードし個人で加工する」との回答が過半数であることは注目される。

一方、「加工」の内容としては、テキストデータをワープロソフトにコピーした上で好きな時に検索できるようにする[12]という程度なのか、さらに高度な加工をしているのかは、不明確で今後の課題とする。

また、「諸本底本情報確認」も回答者の21名が選択している。底本に関しては、日本古典文学大系の冊子体に記述されていることであるが、日本古典文学本文データベースのテキストデータだけでなく、関連情報が辞書の様に使用されていると思われる。

### 6.2 研究への影響

また、「研究への影響」については、「成果が出しやすくなった」と評価されており、用例などのデータ収集の効率を高める道具として使われていることが分かる。研究での「資料利用の変化」に対しては、「冊子体よりもデータベースを利用ようになる」という回答がある一方で、「データベースよりも冊子体を利用ようになる」という回答も前者の半数以上ある点が注目される。

前者の回答については、2通りの解釈が可能である。1つは冊子体資料と比較しての検索やデータ収集の効率を上げるためである。もう1つは、より幅広い資料の探索が可能になる点である。文献[13]では、「国歌大観」のCD-ROMを取り上げて、以下のような記述がある。「これくらい演習の光景を変えたものはない。従来、「国歌大観」を悉皆調査して演習に臨むのは、大体が和歌関係の学生だった。今は、物語関係の学生がいとも気楽に膨大なデータを打ち出してくる。散文研究と和歌情報との気分的ハードルが低くなったことを実感する」。この記述からは、専門外の分野の資料であっても、それが電子化されている場合、その資料を積極的に調査しようとし、その結果冊子体よりもデータベースを積極的に利用することになる、という可能性である。事例Aの回答者は、海苔加工の専門家であり、日本の古典文学を専門に扱う分野の研究者ではない。このような事例からも、データベースは冊子体資料よりも探索へのハードルが低いと言えるのではないだろうか。

### 6.3 冊子体資料との補完関係

後者の回答については、日本古典文学本文データベースを利用した後で、冊子体で確認をしていると考えるのが自然であろう。以下の事例のような回答もあった。

#### 事例 H

「このデータベースから直接研究結果を出

すのはためられる。語彙や用例検索に費やす時間は激減したが、見落としが無いか心配。特に本文が旧大系であるため、結局は再び確認のために検索をしている。」

回答者の日本古典文学本文データベースに関する要望からは、日本古典文学大系の電子化テキスト以外の役割を期待されていることが明らかである。「より多くの作品の収録」は回答者のほぼ全員が希望しており、「論文・図書とのリンク」についても約半数の回答者が希望している。日本古典文学大系の作品は全て収録しているにもかかわらず、前者の要望が出てくる背景として、信頼できる電子化テキストデータを手に入れたいという要望があると考えられる。

現在、インターネット上には、様々な文学作品や歴史史料が公開されている。その他にパッケージメディアとして発売されている資料も多数ある [14]。しかし、インターネット上で公開されるテキストデータの多くは、著作権等との関連から最新の研究成果を反映していないケースも考えられる [12]。このような状況の中で、研究に使用できる精確で信頼できる電子化テキストデータへの期待は高まっている。

## 6.4 信頼できる本文

日本古典文学大系以外の冊子体資料も視野に入れた以下のような要望もあった。

### 事例 I

新大系、新編全集などをはじめ、新しい本文・信頼できる本文をもとにして、データベースの更新を継続していただけるとありがたい。

この要求は日本古典文学本文データベースを、これ以上情報が加わることはない冊子体の日本古典文学大系と同様に閉じた世界であることを前提にすると、分かりにくいものである。しかし、ひとつの作品に複数のテキス

トが存在する古典作品の実態や、WWW のハイパーテキストなどの機能を考慮すれば、当然の要求とも考えられる。前者に対しては、複数の校訂本が出版されている現状がある。また、古典作品の本文に対する研究は蓄積を続けており、国文学関連の論文については「国文学論文目録データベース」もある [15]。しかし、研究対象とする作品の他のテキストや先行研究を知る道具として、日本古典文学本文データベースからの直接リンクがあればより確実である。しかし、他の本文や論文・図書とのリンクを実現するためには、国文学以外のどの研究分野までをリンク対象とするのか、特定の論文や図書をどの作品 (ファイル) にリンクさせるのが適切なのか、など一定の方針を立てなければならない。さらに、リンク対象となる論文や図書の著作権の問題をクリアする必要もあり、直ちに実現するのは難しいと考えられる。しかし、Z39.50・メタデータなどを活用する開発研究はかなりすすんでいる [7]。

## 6.5 本文以外の情報とのリンク

日本古典文学本文データベースに関する要望には、「論文・図書とのリンク」とほぼ同数の回答があった「外字の取り扱い」はさらに複雑な問題である。日本古典文学大系の底本から活字テキストの段階では翻刻者の解釈が、活字テキストから電子化テキストの段階では、文字コードによる制約がかかっている。いくつかの可能性のある文字については、その情報を参照できる機能を求める事例 J の様な利用者の意見もあった。

### 事例 J

「旧字、正字、俗字、異体字の区別無く検索できるようになるとより便利である。例えば「四庫全書」の CD-ROM 検索のように、語彙、単語レベルでの検索ができるようになると使い勝手が上がる。例えば、活用語の活用語尾の違いに対応できるようにする、また

異字同訓の語が同時に引けるようにする。」

## 6.6 その他の要望

その他の要望としては「複数作品 (ファイル) を同時に文字列検索したい」という回答が 11 名あった。この要望は 1 名は国語学の研究者からであったが、後は全て国文学の研究者からであった。ただし、この機能はインターネットブラウザの複数のウィンドウを用いて実現可能であり、その方法も公開されている [16]。しかし、その使い方に気がつかなかったということであろう。

なお、国文学研究も国語学研究も古典作品を対象として用例を確認する点は共通である。自身の研究テーマに応じて、複数の作品から特定の語句を共時的あるいは通時的に収集する。そのため、複数の作品を同時に検索できればより利便性が増す。複数の作品を検索して行われた研究としては、事例 C や事例 K がある。後者の様に、利用者自ら複数の作品を同時に検索できるように工夫をしている場合もある。

### 事例 L

「あるキーワードがどのように使われているかを調べるのに最適である。とりわけ WZ のグローバル検索を使うことによって、古典大系すべての検索が可能になった。」

## 7 おわりに

本調査からは、日本古典文学本文データベースの利用者の評価と様々な要望が明らかになった。利用登録者が約 1400 名であることを考慮すると、利用者全体の傾向として捉えることには無理があるが、日本古典文学本文データベースを高頻度で利用する研究者の意見を集約できたことは意義がある。

本稿のもう 1 つの目的である日本古典文学本文データベースは研究に影響を与えるのかについての考察は難しい。アンケートの回答

からは、研究作業能率の向上や分析できるデータ量の増大は指摘できるが、研究に質的な影響を与えたかどうかは、研究者間の見解に相違がある。「新しい用例の検索が容易になり、注釈に新たな知見が得られた」という国文学者の評価がある一方で、「(研究の) 根幹部への影響はない」という国語学者の言及もあり、この段階で結論を出すことは出来ない。しかし、事例 A としてとりあげた海苔加工の専門家の発言や事例 B の米国の歴史学者の発言のように、日本古典文学本文データベースの存在を前提とした調査が行われていることは事実である。

日本古典文学本文データベースの利用者の多くは国文学、国語学、歴史学の研究者が中心である。これらの異なったいずれの研究分野に対しても有用なデータベースであるためには、何が必要なのだろうか。また、「データベースは、その道の専門家よりも、門外漢や初学者がより有り難みを感じるものである。だからこそ使い方を誤り、なかなか必要とするデータにたどり着けないといったジレンマを感じることもある。データベースは、学部の学生から手軽に使えるものだからこそ、その構築には配慮が必要であるし、また、使い勝手がよくなるように改良を重ねていくべきものだとも思う」[17] という指摘もある。現在の学生は、コンピュータをレポートや論文執筆など日常的に利用しており、その延長で日本古典文学本文データベースの利用をする可能性も高い。このように、学生など古典文学に対して、入門者である利用者に対してはどのようなサービスが適切なのだろうか。

日本古典文学本文データベースを日本古典文学大系の単なる電子版ではなく、研究者にとってより有用なデータベースの一つとするために今後の課題としたい。

## 謝辞

本稿は 2002 年度情報知識学会研究報告会での発表に基づいたものである。アンケート調



査に協力していただいた方々および貴重なアドバイスを下さった方々、調査を許可し、研究を指導していただいた国文学研究資料館および安永尚志先生に心より感謝申し上げます。

## 参考文献

- [1] Genicot, Leopold(大嶋誠 [ほか] 訳):「歴史学の伝統と革新:ベルギー中世史学による寄与」, 福岡, 九州大学出版会, 248p, 1984.
- [2] 神立孝一:「近年人文研究におけるコンピュータ利用をめぐって:近刊『電脳国文学』(漢字文献情報処理研究会編)を中心に」, 弘前大学国史研究, No.110, pp.56-61, 2001.
- [3] 村主朋英;山西史子;松井美紀:「歴史学におけるインターネット情報源の構造と利用状況」, 日本図書館学会研究大会発表要項, 第47回, 日本図書館情報学会, 大阪, 1999, pp.9-12.
- [4] 伊井春樹:「古典文学の情報検索と今後のあり方:源氏物語データベースを中心に」, 文学・語文, No.171, pp.29-33, 2001.
- [5] 近藤泰弘:「コンピュータによる文学語学研究にできること:古典語の「内省」を求めて」, 文学・語文, No.171, pp.34-49, 2001.
- [6] 安永尚志:「国文学研究とコンピュータ」, 東京, 勉誠社, 500p, 1998
- [7] 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:国文学データ共有のための標準化」, 情報知識学会誌, Vol.11, No.4, pp.17-34, 2002.
- [8] [http://base3.nijl.ac.jp/Rhonmon/hon\\_news.html](http://base3.nijl.ac.jp/Rhonmon/hon_news.html) (アクセス確認 2002/12/19)
- [9] [http://base3.nijl.ac.jp/Rhonmon/hon\\_annai.html](http://base3.nijl.ac.jp/Rhonmon/hon_annai.html) (アクセス確認 2002/12/19)
- [10] 伊藤鉄也:「メディア時評:日本古典文学本文データベースの現状と今後」, 文学, Vol.1, No.3, pp.162-164, 2000.

- [11] 安永尚志:「日本文学研究とコンピュータ」, 文学, Vol.1, No.4, pp.85-93, 2000.
- [12] 漢字文献情報処理研究会:「電脳国文学:インターネットで広がる古典の世界」, 東京, 好文出版, 286p, 2000.
- [13] 後藤祥子:「メディア時評 つまみ食い」, 文学, Vol.1, No.6, pp.226-228, 2000.
- [14] <http://www.konan-wu.ac.jp/~kikuchi/> (アクセス確認 2002/12/19)
- [15] <http://www.nijl.ac.jp/bunseki/index.html> (アクセス確認 2002/12/19)
- [16] [http://base3.nijl.ac.jp/Rhonmon/hon\\_qa.html](http://base3.nijl.ac.jp/Rhonmon/hon_qa.html)
- [17] 川添房江:「デジタル情報化時代の文学研究と大学教育」日本文学, Vol.50, No.4, pp.1-15, 2001.

## 付録: アンケート項目とその回答数

問1 「日本古典文学本文データベース」の主な利用目的は何でしょうか。

a: 研究利用 (86), b: 教育利用 (6), c: その他 (1), 回答者数 = 93

問2 「日本古典文学本文データベース」の主な利用方法は何でしょうか。

a: 一次資料(原資料)として (29), b: 二次資料(レファレンスツール)として (65), c: その他 (0), 回答者数 = 94

問3 データベースの機能の中で主に何を使っていますか。(複数回答可)

a: 本文全文表示・ダウンロード (55), b: 語彙, 文字列検索 (33), c: 文字頻度分析 (1), d: 作品や底本情報の確認 (0), e: 特に決まっていない (4), 回答者数 = 93

問4 データベースを研究・教育に使用する際にどのような使い方をされましたか。(複数回答可)

a: 文字列検索をして用例を集めた (73), b: 特定の場所を読むために文字列検索をした (50),

c: 本文をダウンロードして自分で加工した (51), d: 文字頻度分析を使用した (10), e: 諸本や底本のデータを確認した (21), f: 試用段階で研究には使っていない (4), g: その他 (0), 回答者数 = 94

問5 「日本古典本文データベース」は研究・教育に役立ちましたか。

a: 役に立った (79), b: 普通 (15), c: その他 (0), 回答者数 = 94

具体的な成果の記述の記入者 (34名)

問6 「日本古典文学本文データベース」の利用により御自分の研究動向に変化はありましたか。(複数回答可)

a: 研究の成果が出しやすくなった (37), b: 今までとは異なったアプローチの研究ができるようになった (27), c: 研究にコンピュータを前よりも頻繁に使うようになった (40), d: 変化は無い (11), e: その他 (3), 回答者数 = 94

問7 データベース全般の使用により, 御自分の研究方法(冊子体などの資料利用)に変化はあると思われますか。

a: あると思う(冊子体よりデータベースを使うようになる)(52), b: あると思う(今までよりも冊子体を使うようになる)(12), c: 影響はないと思う (28), 回答者数 = 92

問8 今後「日本古典文学本文データベース」に期待することはありますか。(複数回答可)

a: より多くの作品を収録してほしい (89), b: 機能を充実させて欲しい (25), c: 本文以外のデータ(論文・図書・原資料の画像等)とリンクさせて欲しい (46), d: 外字が扱えるようにして欲しい (45), e: 特になし (1), f: その他 (13), 回答者数 = 94

具体的な要望の記入者 (30名)

問9 同種の本文データベース(青空文庫など)でよく利用されるものがあれば, 下にご記入をお願いいたします。記入者 (41名)

問10 その他何かご意見があれば以下にお願いします。記入者 (37名)

(2003年1月8日受付)

(2003年2月20日採録)

## 2002 年情報知識学会月例懇話会報告

月例懇話会担当 平田 周

情報知識学会では、会員たちが、気楽に議論し、交流を図ることを目的に月例懇話会を開催しております。2002年は、2月、3月、6月、7月、9月、10月、11月と7回開催しました。その都度、開催案内とともにメールで会合報告をお送りしておりますが、ここにまとめておきました(10月度は出席者がなく、議事録はなし)。今年もぜひ多数の方々が集まって、会を盛り上げてくださるようお願いいたします。

## &lt; 2 月度月例懇話会 &gt;

開催日 2002年2月5日(火曜日)午後6時半～9時  
講師 大野 照夫氏(TBS 番組考査部部長, 情報知識学会会員)  
話題 テレビ局から見た情報問題の諸相  
参会者 岸本吉浩, 五所吉哉, 中山亮一, 平田周

TBSのディレクターとして多忙な日々を送っておられる大野照夫氏に講師をお願いし、メディアから見たさまざまな問題について話してもらった。同じメディアで活躍されている、東洋経済新報社の岸本吉浩氏が初参加下さり、盛り上がった。

大野氏は、昨年、メディアリテラシーをテーマに、民放連・NHKの共同企画番組「テレビ観察プロジェクト 21世紀のテレビはこうなってほしい!! 子ども達の願い」のプロジェクトをプロデュースされた。高知市立旭東小学校の6年生を集め、テレビの観察日記をつけたり、放送局を訪ねて番組ができるまでの過程を見学したり、街頭でアンケートをとる、ビデオをつくる、グループごとに未来のテレビに関する提言を発表するといった試みを番組にまとめたもの。ただ漠然とテレビを見るのではなく、なぜだろうという問題意識をもってテレビと対面する時、そこに子ども達が新しい世界を発見する様子が明らかにされた。情報化時代、情報を見極め活用する能力、情報を読み解く力が求められる。ここに将来のテレビの新たな役割を考えるきっかけとなるものが見いだされたという。

話は、さらにテレビと政治の問題などにも及び、放送業界が直面するさまざまな問題について、その分野にいそしむ人しかわからないことを興味深く聞くことができた。

## &lt; 3 月度月例懇話会 &gt;

開催日 2002年3月5日(火曜日)午後6時半～9時  
講師 中山 亮一氏(情報知識学会会員, 電子情報通信学会用語特別専門委員, 日本ポストプロダクション協会広報委員ほか)  
話題 専門用語研究の必要性 オンライン学術用語集のことなど  
参会者 井上孝, 太田泰弘, 土本茂, 仲本秀四郎, 平田周

中山亮一氏がお話し下さるということで、当学会専門用語部会の面々が集まれ、同部会の小集会のようなになった。とくに仲本先生がご参加下さったので、話題はあっという間に盛り上がった。こういう同じテーマに関心のある方々の集まりに、月例懇話会を利用されるのもよいと感じた。

中山亮一氏は、もともとテレビ用語学会の用語委員会のメンバーとして用語事典の編纂などに携われ、専門用語の研究が、まずは学問的なことよりも、処理など実用面での問題の重要性を認識されたという。とくに通信学会やテレビ学会、あるいは電気学会など、学会ごとに用語の使

い方が異なっているなど不便この上ない。文部省は、昭和 22 年から用語に基準を作る試みを始めるなど、それなりに関心は高く、各専門分野学術用語集を発行した。しかし、これらは、各学会からの協力が中心だったため、学会ごとに違う用語の調整や統一まではできず、著作権の問題もあって、勝手な変更もできないという事情にあった。

現在、国立情報学研究所が「オンライン学術用語集」をつくり、同研究所のウェブサイトで利用することができる (<http://sciterm.nii.ac.jp/>)。対象となる分野は、遺伝学編、海洋学編、気象学編、キリスト教学編、計測工学編、言語学編、原子力工学編、建築学編、航空工学編、採鉱や金学編、地震学編、心理学編、数学編、船舶工学編、地学編、地理学編、天文学編、図書館情報学編、土木工学編、物理学編、分光学編、論理学編にとどまっている。各関連学会からの利用承認を得るのに時間がかかっているようだ。

このような用語集の編纂や利用で問題になるのは、専門用語が未だ不統一だという点にある。たとえば、コンピュータかコンピューターかで検索できるものが異なる可能性がある。そういう意味でも、専門用語の研究は必要不可欠であるが、学問的方向からよりも、現場的な言葉に対する関心から進めていくことが必要であろうと考える。

#### < 6 月度月例懇話会 >

開催日 2002 年 6 月 18 日 (火曜日) 午後 6 時半～9 時

講師 迫 猛氏 (三笠書房取締役書籍編集部長)

話 題 出版界の最近のことなど

参会者 安藤梢, 岩渕幸雄, 太田泰弘, 岡伸人, 衣川純一, 五所吉哉, 佐藤純一,  
高橋靖明, 根岸正光, 平田周, 月見里禮次郎

わが国の出版業界は、1995 年以後、不景気に見舞われている。統計を見ても、新書籍発行部数は、1997 年の 15 億 7,354 万冊をピークに以後減り続け、2001 年は 13 億 8,578 万冊に落ちている。一方、発行点数は、1995 年の 58,310 点から 2001 年には 71,073 点へと増加の一途をたどっている。

出版界が低迷に苦しむなか、三笠書房は元気がある。その理由を、迫猛編集部長は、たまたま企画したものがヒットしただけではあるが、売れそうな本を狙って編集をすることが決め手になっていると説明する。ではどのような本がよく売れるのか。具体的ニーズ、啓蒙的なものに読者は関心が高い。残念ながら、多数の読者が買う本は、キーワードとして「女」「食」「金」(ビジネス)ということになるというのが業界の常識である。本が売れないといわれるが、100 万部以上売れるものもけっこうある。

若い人たちの情報や時間の使い方に関する価値観が変化したことが大きいという。とりわけ、我慢することを知らない、本を買わなくなったなどが目立つ。会に出席していた大学院生から、本を読まなくなった理由に、(1) 家に昔のように文学全集が並んでいるといった「本のある生活」がないこと、(2) あまりに本が多く、どれを読んでもいいかわからない、という意見が出された。本を読むのにかかる時間は昔と変わらないのに、現代生活の時間リズムはものすごく早くなっていてバランスがとれないということも関係しているのではないかという意見も他の出席者から出た。

野球の野村克也氏に迫氏がインタビューした時、人材に話が及び、プロ野球選手で伸びるのは、出来のよい優等生でかつ不真面目な、いちばん多いのは、劣等生で真面目な選手だと語っていたのが印象に残っていると話す。編集の世界でも、このことはあてはまるように思う。

< 7 月度月例懇話会 >

開催日 2002 年 7 月 18 日 (木曜日) 午後 6 時半～9 時  
講 師 名和 小太郎氏 (当学会理事, 国際大学 GLOCOM 客員教授)  
話 題 ゲノム特許 - 研究・ビジネス・倫理  
参会者 岩渕幸雄, 太田泰弘, 岡伸人, 川俣修壽, 菊田昌弘, 白鳥裕, 杉浦裕美,  
高橋仁一, 平田周, 細野公男

ヒトゲノム解読以来, バイオテクノロジーはさらに未知の世界を解き明かしつつある。ビジネスや研究に携わる人たちには, 次々となされる発見や発明に特許権を得ようとするのは当然であるが, その一方で, 生命や医療などにかかわる技術を一部の者の手に独占されているのかという, 自然保護あるいは倫理的方面からの批判もある。

生物特許についての問題は, 自然の産物は特許の対象としないという特許本来の考えと抵触する恐れが強い。アメリカで植物特許法が成立したのは, 1930 年のことだが, 画期的な出来事は, 1980 年にチャクラバーティの人工微生物特許を米最高裁が認める判決を行ったことであった。同年には, コーエン, ボイヤーの遺伝子組み換え特許も成立しており, 人間の創造したものすべてに, 特許が認められる方向が鮮明になった。

こうしたアメリカの動きに対し, ヨーロッパは慎重である。1998 年に制定された EU バイオテクノロジー発明保護指令は, 人体とその要素の単なる発見や公序良俗に関するものは不特許とすることを明らかにしている。

とりわけ DNA に関する特許は, 自然の産物ではないかとする疑問に対し, 抽出・精製は特許の対象となりうると反論するなど論争を生んでいる。

また, そのデータを特定の営利企業が所有し, 使用に制限を加えることの不当性を主張する科学アカデミー, 王立協会の意見などもある。UNESCO は, ヒトゲノムに関する人権宣言を 1997 年に発表した。「自然状態にあるヒトゲノムは財政的利益を生じさせてはならない (第 4 条), ヒトゲノムに関し, 生物学, 遺伝学, 医学の進歩から得られた利益は, 個人の尊厳と人権に関する十分な配慮を払って, すべての人が利用しうるようにしなければならない (第 12 条 a)」と宣言している。

参考図書 : 名和小太郎著「ゲノム情報はだれのものか」岩波書店 2002  
名和小太郎著「学術情報と知的所有権」東京大学出版会 2002

< 9 月度月例懇話会 >

開催日 2002 年 9 月 18 日 (水曜日) 午後 6 時半～9 時  
講 師 仲本 秀四郎氏 (元日本原子力研究所技術情報部長, 茨城大・東京学芸大講師など  
歴任)  
話 題 知識とターミノロジー  
参会者 井上孝, 太田泰弘, 長田孝治, 笹森勝之助, 佐藤純一, 高橋靖明, 田良鳥哲,  
名和小太郎, 春山暁美, 平田周, 細野公男

知識とターミノロジー

1. 概念を知識の単位ととらえ, 概念を記述したものが定義, 定義の短縮形もしくはラベルが用語という立場からターミノロジーを考えていると前提して, ターミノロジーの歴史, 他分野との関係を解説した。知識は概念の規則性をもつ集合であって, その体系性が知識構造, 論理ルールに従って経験から得られるというアリストテレスの説を紹介した。

2. 紀元前7世紀の法律用語の対訳表を嚆矢とし、中世を脱して、言語関係新活動の台頭から組織化され、標準語運動などの言語政策が始まり、20世紀に入って、概括的・体系的な方法論を E. Wuester が確立した。

動機としては、(1) 言語産業の勃興、(2) 流通制度と標準化、(3) データの作成・蓄積・検索、(4) 科学技術文書の翻訳、(5) 技術移転での必要性、(6) 政府の干渉、(7) 言語政策での保護主義などが挙げられた。

アプローチとして、主題志向、哲学志向、言語志向の3つが特徴的で、(1) ウィーン学派の一般理論、(2) プラハ学派の機能言語学の視点、(3) ソビエト学派の規制志向、(4) カナダは2言語国家で、ケベックのフランス語活動とかかわりあっているなどが紹介された。

3. 各分野との関係では、(1) 辞書学が意義論的 (semasiological) であるに対し、ターミノロジーは名義論的 (onomasiological) であるのが違いであり、通時性の欠如が特徴、(2) 言語学とは方法論・表示法・造語などでターミノロジーの自立性が唱えられ、ソシュールの示唆が重要である、(3) 論理学とは概念と概念関係で基本的考察を借用、(4) オントロジーは時間的空間的位置づけで深く関わりあっている。(5) 情報学とは情報表示と分類で親密な関係にあり、(7) ターミノロジー作業は主題専門家によって遂行される。

シソーラス：Roget のシソーラスは、概念志向・組織化している点で、構想がターミノロジー的である。一方、検食用シソーラスはシステム依存・高改訂頻度・無定義・アルファベット配列、全体系を示していないなどで、ターミノロジーと異なっていると指摘があった。

分類との関係：対象に即して帰納的に分類することはあまりなく、演繹的な概念体系を借用しているため、アプローチの逆方向性からターミノロジーの概念体系とズレを生じている。概念体系による分類では、対象に対応しているかの検定が必要と指摘した。

話の過程で、アルファベット配列が16世紀からと比較的新しく、知識の一般への普及に伴い、アクセスの容易さが動機になっていると紹介された。また、今回、実務には触れる余裕がなかったが、Technolexicon の失敗(1900年)から、計画にあたっては、野心の抑制がターミノロジー作業での精神性であると言及があった。

#### < 11 月度月例懇話会 >

開催日 2002 年 11 月 18 日 (月曜日) 午後 6 時半～9 時

講師 岩田 修一氏 (当学会理事、東京大学教授・人工物工学研究センター)

話題 データ科学と CODATA の社会的役割

参会者 岡伸人、五所吉哉、佐藤純一、杉浦裕美、時実象一、名和小太郎、馬場哲也、平田周、山本毅雄

当日ご出席いただいた国立情報学研究所の山本毅雄先生は、講師をつとめて下さった岩田修一先生の恩師、そしてその山本先生は、当学会の藤原鎮男会長から70年代初め、当時まだめづらしかった環境問題について講義を受けられたという関係から、当時のいろいろなことについて、話に花が咲いた。

ところで、岩田先生は、世界的に知られる CODATA (Committee on Data for Science and Technology) の会長になられ、CODATA の歴史的経緯、その役割などについて話された。CODATA の目的は、基礎物理常数の設定のような地味な仕事から始まり、コンピュータ利用の発展とともに、データの収集、データベースの管理から、データの品質、利用の容易さなどの問題へと広がっていった。一言でいえば、CODATA の役割は、データを中心にして、世界的な広がりの中で知識の共有化を進めるところに任務がある。

現在は、データベースとして単に情報を提供するというところから、それらのデータを解釈するモデルも提供する必要性が認識されるに至っている。情報のデジタル化は、ものに対する臨場感がない、マニュアル的感觉であり、リアリティの本質をつかめないといった欠点を持つ。

今後は、より人間的な要素が重視され、情報にインタラクティブな機能を持たせることが重要になると考える。

なお、CODATA の日本支部は、情報知識学会であり、岩田先生が会長に就任されたいま、学会としてもっと CODATA の活動に関心を持ち、応援しなければならないと感じた。科学技術の基礎データの管理というのではなじみにくいですが、そうした重要な情報を世界的に知識として共有し、その使い方をより便利にし、さらには岩田先生が主張されるように、人間的側面まで重視していくということであれば、当学会の研究や活動にも密接なかわりがあるに違いない。

## 「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、編集委員会を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も編集を行う。

## 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ない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（暫定版）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] 藤原譲: 「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 数式は印刷に便利の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，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

最近の日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、かつ当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

## 平成 15～16 年度役員選出について（中間報告）

標記の役員選出手続きに関連し、理事会として推薦する役員候補者リストを今号で発表する予定ですが、諸般の事情により来月、封書にてお知らせします。

今後の日程は以下のようになります。

- (1) 総会通知とともに投票方法の説明を正会員全員に郵送し、投票を求める。同封の投票用紙（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。

## 《編集後記》

13巻1号をお届けします。2月末発行予定でしたが、1ヶ月遅れとなってしまいました。誠に申し訳ありません。今年(13巻)から、本会誌は暦年での発行となります。その第1号です。なお、14巻以降の第1号は1月発行を目指します。また、1号(1月)、2号(4月)、3号特集号(7月)、4号(10月)となります。

編集委員会は、年4回の定期発行を最重要課題として取り組んでいますが、依然として投稿論文が少ないという課題があります。現在、会員数も300名強ですので、母集団自体が小さいからと思われれます。まず、会員諸氏の周辺から、新たな会員の獲得をお願いします。加えて、是非ご投稿をお願いします。

本号は研究論文2本、論談1本、調査報告1本の論文を掲載しました。各論文が対象とする分野も生物学、システム学、人文学に渡り、幅広いものとなっています。本学会の目標とする多様な分野における情報・知識に関わる課題に応えるものです。各論文の専門性は高く、新奇性、有用性など評価も高いのですが、専門外の読者にも興味を持っていただけるよう論旨、書き方などに工夫がなされています。是非、専門外の方々もお読みいただきたいと思います。できれば、質問やコメントなど、また読後感をお寄せください。

2003年3月 安永 尚志

---

---

### 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  |

---

---

□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@muj.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.13, No.1 2003年3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/>

---

---

