

Journal of Japan Society of Information and Knowledge

情報知識学会誌

Vol. 16, No. 2 (May, 2006)

~~~~~ 目 次 ~~~~~

## 特集 第14回(2006年度)年次大会(研究報告会&総会)

科学研究費データベースによる採択と研究成果のパターン分析

.....西澤正己、根岸正光、柴山盛生、孫 媛、渡邊恵子、野村浩康、光田好孝、前田正史... 1

日本の大学と企業の研究リンケージ—NCR-JとCJPデータベースによる比較

.....孫 媛、根岸正光、西澤正己、渡邊恵子...7

ISIデータベースにおける発表論文数、被引用回数に基づく「大学ランキング」作成の方法と問題点

..... 根岸正光...13

最適解の付与に基づく重み付けの自動変更..... 原田隆史...19

新たな文脈を生成するデジタルコンテンツ構成法.....宇陀則彦、松村 敦、村田典子...23

問題解決を目的としたWWW情報探索行動の特徴..... 相良佳弘...29

ユーザの検索要求に基づいた興味関心の定量的評価.....堀 幸雄、中山 堯...33

Web上における『法政経済社会論文総覧』の検索閲覧システムの構築..... 古隅弘樹、周防節雄...39

科学技術文献検索システムにおける大規模用語辞書の活用について.....甲田 彰、森田歌子...45

ミュージアム資料情報構造化モデル応用の検討..... 斎藤伸雄...49

脳の情報処理と問題解決..... 福永征夫...53

インターネットを利用した遠隔健康管理システムの試み.....里村宏章、田中猛彦、吉廣卓哉、中川 優...59

テキスト入力によるキーストロークダイナミクス.....佐村敏治、西村治彦...63

建築とドキュメントについての考察..... 山本隆彦...69

歴史から洞察される判断基準としての観点.....安平哲太郎、佐脇政隆...75

熊本県医師家塾調明治6年による医療教育と使用教科書(35事例)..... 安澤 秀一...81

追悼:初代会長 米田幸夫先生を悼む..... 87

投稿規程と執筆要領..... 88

事務局からのお知らせ..... 96

~~~~~



情報知識学会

<http://www.jsik.jp/>

※年次大会には本会誌をご持参ください。

第14回（2006年度）年次大会（研究報告会&総会）プログラム

日程：2006年5月26日（金）、27日（土）

会場：近畿大学B館（法科大学院・経済学部棟）10階MM会議室

5月26日（金）

13：00 受付開始

13：30～13：40 会長（細野公男）開会挨拶

13：40～14：20 総会

14：20～14：30 実行委員長（田窪直規）事務連絡

14：30～14：40 休憩

14：40～15：55 発表1（計量文献学関連研究） 座長：中川 優（和歌山大学）

発表3件 持ち時間各24分（発表18分、質疑6分、交代時間1分）

1. 西澤正己（国立情報学研究所）ほか、科学研究費データベースによる採択と研究成果のパターン分析。
2. 孫媛（国立情報学研究所）ほか、日本の大学と企業の研究リンケージ：NCR-JとCJPデータベースによる比較。
3. 根岸正光（国立情報学研究所）. ISIデータベースにおける発表論文数、被引用回数に基づく『大学ランキング』作成の方法と問題点。

15：55～16：10 休憩

16：10～17：00 発表2（情報組織化関連研究1） 座長：長塚 隆（鶴見大学）

発表3件 持ち時間各24分（発表18分、質疑6分、交代時間1分）

4. 原田隆史（慶応義塾大学）. 最適解の付与に基づく重み付けの自動変更。
5. 宇陀則彦（筑波大学）ほか、新たな文脈を生成するデジタルコンテンツ構成法。

17：20～18：50 懇親会（於：近畿大学11月ホール地下1F食堂（Cafeteria November））

5月27日(土)

9:45~11:00 発表3(情報検索関連研究) 座長:細野公男(慶應義塾大学(名誉教授))

発表3件 持ち時間各24分(発表18分、質疑6分、交代時間1分)

6. 相良佳弘(聖徳大学). 問題解決を目的としたWWW情報探索行動の特徴.
7. 堀幸雄(香川大学)ほか. ユーザの検索要求に基づいた興味関心の定量的評価法.
8. 古隅弘樹(兵庫県立大学)ほか. Web上における『法政経済社会論文総覧』の検索閲覧システムの構築.

11:00~11:15 休憩

11:15~12:30 発表4(情報組織化関連研究2) 座長:山本 昭(愛知大学)

発表3件 持ち時間各24分(発表18分、質疑6分、交代時間1分)

9. 甲田彰(科学技術振興機構)ほか. 科学技術文献検索システムにおける大規模用語辞書の活用について.
10. 斎藤伸雄(凸版印刷). ミュージアム資料情報構造化モデル応用の検討.
11. 福永征夫(アブダクション研究会). 脳の情報処理と問題解決.

12:30~13:45 昼休み

13:45~15:00 発表5(事例的研究) 座長:平田 周(東京工科大学)

発表3件 持ち時間各24分(発表18分、質疑6分、交代時間1分)

12. 里村宏章(和歌山大学)ほか. インターネットを利用した遠隔健康管理システムの試み.
13. 佐村敏治(明石工業高等専門学校)ほか. テキスト入力によるキーストロークダイナミックス.
14. 山本隆彦. 建築とドキュメントについての考察(仮題).

15:00~15:15 休憩

15:15~16:05 発表6(歴史関連研究) 座長:吉野敬子(三菱ウェルファーマ)

発表2件 持ち時間各24分(発表18分、質疑6分、交代時間1分)

15. 安平哲太郎(産業技術総合研究所)ほか. 歴史の指し示す判断基準としての観点.
16. 安澤秀一.(国文学研究資料館(名誉教授))熊本県医師家塾調明治6年による医療教育と使用教科書(35事例).

16:05~16:15 副会長(安永尚志)閉会挨拶

科学研究費データベースによる採択と研究成果のパターン分析

○西澤 正己、根岸 正光、柴山 盛生、孫 媛、渡邊 恵子(国立情報学研究所)
野村 浩康(東京電機大学)、光田 好孝、前田 正史(東大・生研)

A Pattern analysis of award of Grant-in-Aid for Scientific Research with Research Outputs

Masaki NISHIZAWA, Masamitsu NEGISHI, Morio SHIBAYAMA, Yuan SUN,
Hiroyasu NOMURA, Yoshitaka MITSUDA, Masashi MAEDA

The system of Grants-in-Aid for Scientific Research from Ministry of Education, Culture, Sports, Science and Technology of Japan is one of the oldest ones, which is the funding system for researchers belonging to universities and institutes in Japan. We report the adoption patterns of the Grant-in-Aid from 1985 to 2002 about the researchers who were awarded the large-scale funds in recent years. Moreover, we mention about the correlation between adoption pattern and research output pattern experimentally.

要旨

筆者らは、科研費採択およびその遡及、成果概要データベースを利用し、近年の大型研究費の採択者、学術賞受賞者等について、1985年から2003年までの科研費採択状況を分析した。さらに、同期間の研究成果を JSTPlus、JMEDPlus、雑誌記事索引等のデータベースから抽出しそのパターンと相関について試験的分析する

はじめに

近年、研究評価の浸透と共に研究資金の配分における重点化の傾向が一層強まりつつある。その一方で、ノーベル賞受賞者等から、基礎研究の重要性と、その資金調達困難化による基礎研究の将来に対する危惧が表明されている。

文部科学省科学研究費補助金(科研費)は古くから基礎研究の下支え、底上げの役目を果たしてきたと言われている。現在では、科研費以外にも多くの競争的資金が配分されているものの、研究費総額では、科研費が約半額を占めている。この他、学術研究の基礎研究に対する資金としては各大学、研究所の校費等が考えられるが、これら基礎的資金と、実用化された革新的技術や成功を収めた画期的研究等との関係を明らかにし、資金配分の重点化が、今後の科学技術の発展にとって真に有効な手段であるのかどうかを定量的、実証的に検証することは重要である。

本研究においては、これまでに実用化された技術や成功を収めた研究等でトピカルなものを抽出し、それらのトピックに対して多面的な調査を実施することにより、それらトピックに対する萌芽的な研究を特定し、その後における研究の推移を追跡する。これにより、これまでの研究費を取り巻く状況中で、どのような研究資金に依拠して始まった研究が今日の成功を収めているかの経緯を明らかにし、中長期的視点での研究予算の効率的配分方式等について考察していく予定である。また、研究補助金による研究の成果をビブリオメトリックス的手法により分析し、研究費とそのアウトプットの関係性を定量的に導出すること等を目的とする。

われわれはこれまでに、特別推進研究(COE)の代表者等39名を選び出し、1985年から2002年までの科学研究費補助金の採択状況を調べ、年度ごとの採択状況にいくつかのパターンがあることを見つけた[1]。その後、対象者を紫綬褒章、日本学士院賞、文化勲章、恩寵賞等の受賞者に広げ、科学研究費の採択状況とその研究成果としての学術論文や著書を調べている。

今回は、学術賞受賞者の採択パターンを分野別に調べ、前回との比較を行った。また、試験的に前回調査した特別推進研究(COE)採択者について、JSTPlus、JMEDPlus、雑誌記事索引等のデータベースから対象者の研究成果を抽出し採択パターンとの比較をおこなった。

1. 科学研究費補助金データベース

国立情報学研究所では、科学研究費補助金の採択課題の速報を収録した「科学研究費補助金採択課題データベース(KAKENK)」[2]およびその成果概要を収録したデータベースである「科学研究費補助金成果概要データベース(KAKEN)」[2] (注)が作成されている。本報告の分析時点で、KAKENKにおいては1985年度から最新の2005年度(注)までの採択課題、KAKENにおいては1985年度から2003年度(注)までの成果概要が収録されている。KAKENKデータベースには研究代表者、採択課題名、研究種目、分科細目(大型研究費、奨励研究を除く)、新規・継続の別、配分額等が収録されており、KAKENについては配分額を除くKAKENKの情報に加え、研究分担者、成果概要、キーワード、発表文献リスト等が収録される。なお、KAKENデータベースは成果概要の報告書を収録するため、採択後少なくとも約1年半後の課題が収録されていることになる。今回の分析には1985年から2003年度までのKAKENKデータベース、および1985年から2003年度のKAKENデータベースを分析に使用した。

(注)現在この二つのデータベースはユーザーインターフェースが統合されKAKENデータベースとして公開されている。さらにKAKEN K(採択課題)は遡及入力が進み、3月31日からは1965年以降のデータが公開されている。

2. データ作成手法と総論

今回は、平成7年から16年までの紫綬褒章、文化功労者賞、日本学士院賞、恩寵賞、エジンバラ賞の受賞者を選び出し、そのうち70歳以下168名を対象者とした。これらの研究者について代表者および分担者となっている採択課題の抽出のため、まずKAKENデータベースから該当対象者の研究者番号により代表者(1)と分担者(2)の採択課題を抽出した。これらの採択課題中1985年以降のKAKENKデータベースに収録されているものについては、課題番号からこのデータベースより補助金の配分額を抽出した。この過程により対象研究者の1985年度から2003年度までの代表者、および分担者となった対象者143名について科学研究費補助金の採択課題と配分額のデータを作成した。作成にあたっては、データベース内での対象者の名前の漢字の表記ゆれ、研究者番号の間違い同姓同名問題等に注意を払って行った。なお、今回の調査には前回の調査に含まれていた研究者は8名あった。

表1-(1)は、今回対象とし、科学研究費データベースに採択課題が1件以上あった143名についての年齢構成であり、表1-(2)は調査期間中(1985年から2003年)における分野別の代表、及び分担採択課題の一人当たり合計配分額(千円)の平均と採択件数の総数である。これをみてわかるように、生物系の医歯薬学、生物学が極めて高く、続いて理工系の工学、化学と続いている。人文系、社会系は配

分類が他に比べてかなり低いことがわかる。尚、課題数については総数であるので注意していただきたい。

表1: (1)対象学術賞受賞者中科研費採択者(1985-2003年)の分野別年齢構成と、(2)配分額の平均および採択課題数の合計

分野	合計	55未満	55-59	60-64	65-70	代表配分額の 平均(千円)	代表課題 数合計	分担配分額の 平均(千円)	分担課題 数合計	人数
社会	11	0	1	5	5	63936.4	111	57436.4	103	11
人文	13	0	4	4	5	32123.1	83	62950.0	113	13
生物-医歯薬学	26	1	4	4	17	667712.7	687	981980.8	665	26
生物-生物学	15	1	3	5	6	707820.0	331	1018146.7	502	15
生物-農学	15	0	0	2	13	131340.0	287	129646.7	272	15
複合	1	0	1	0	0	309000.0	26	290200.0	29	1
理工-化学	16	0	1	4	11	392581.3	359	231550.0	224	16
理工-工学	18	2	3	5	8	541947.2	362	393577.8	403	18
理工-数物	28	1	3	9	15	193875.0	437	184477.8	488	28
全体	143	5	20	38	80	369529.2	2683	422680.6	2799	143

3. 年次別採択金額のパターン分析

ここではまだ統計数が少ないためと、1985年以前の採択課題遡及データが現時点でまだ使用できなかったために、前回と同様にパターンフィッティング等の手法は使っていない。また、高齢の方の調査を正確に行うには1985年以前のデータが必要なため、ここではさらに2005年の時点で65歳未満の対象者にしぼった。これらの62名の対象者に対して、人文系、社会系、理工系(化学、工学、数物)3分野、生物系(医歯薬学、生物、農学)3分野の8分野別に前回のパターンとの相関をみている(複合系1名は分析対象から外した)。パターンは基本的に代表者として受けた科研費の配分額を元に分析している。また、研究分担者としての額は、代表者が受けた額をそのまま統計しており、分担者数等の補正もおこなっていない。基盤研究の一部で、研究(1)の形態をとっているもの以外、ほとんどが分担金を受け取っていない場合が多く、研究分担の度合いもデータからはわからないので、分担者としての額を見るときにはこの事を念頭に入れ注意をして見ていただきたい。ここでも、分類の参考程度にするにとどめている。グラフについては縦軸を対数スケールとしているのでこの点も注意してほしい。

3.1 前回の分析による採択パターン

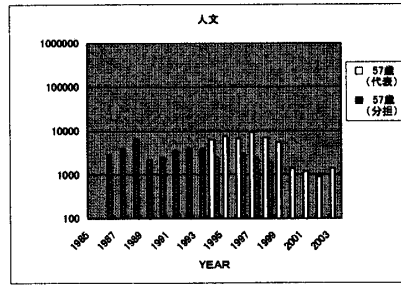
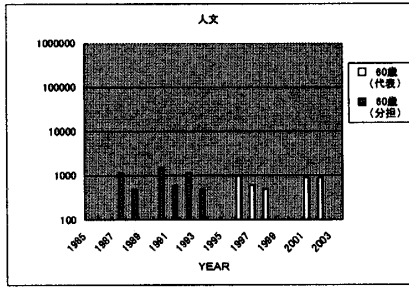
前回の分析では、特別推進(COE)33名と著名研究者6名について、採択金額の年次変化を以下の5つのパターンに分類した。

- (1) 理論研究タイプ: 小型の採択課題が少数見られるだけで、理論研究者に多い。(2名)
- (2) 積み上げタイプ: 小額の補助金から徐々に中型、大型と移っていくタイプ。(16名)
- (3) 高額維持タイプ: 高額を維持。若い頃のデータがない可能性もある。(11名)
- (4) 急出現タイプ: 海外からの、帰国と共に大型研究費の代表者となる場合が多い。(5名)
- (5) 世代交代タイプ: 小型の代表の後、大型の分担を経て大型の代表者になる。(5名)

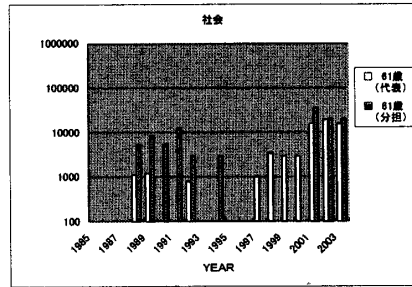
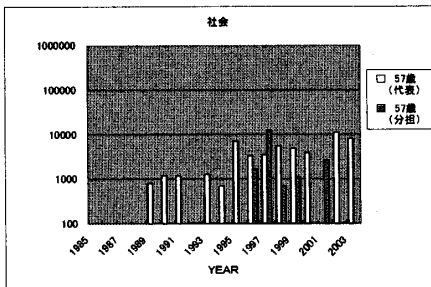
今回は、分野別の採択パターンとこれら5つのパターンとの比較をおこなう。

3.2 分野別の採択パターン

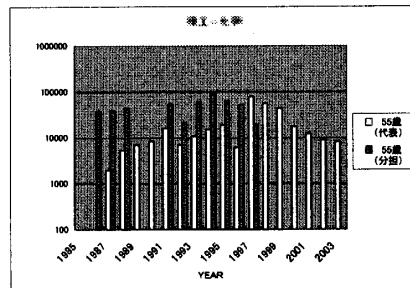
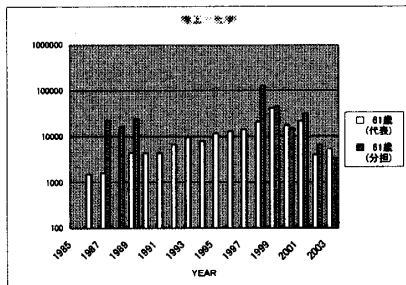
(A) 人文系: 人文系については全般的に他分野と比較して科研費に依存する傾向は小さい。人文系では対象者中科研費の受給割合は68%と低い割合となっている。よって、全般的な統一傾向はなく、少額が多いことや、長期間継続して研究費を受けている研究者が少ないといった傾向などが見られる程度である。中には史学関係で分担を含めると長期にわたって受給している人もいる。



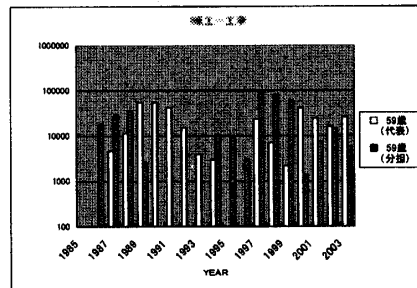
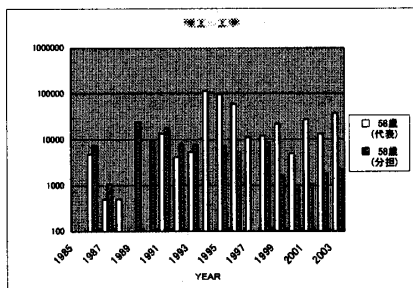
(B) 社会系: 社会系に関しても人文系と同様な傾向であるが、やや金額が大きくなっている。また、中には積み上げタイプに近いタイプも見られる。科研費の受給割合は人文系とほぼ等しく69%であった。



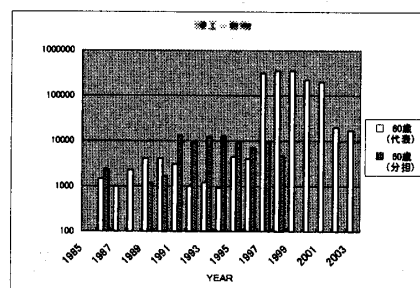
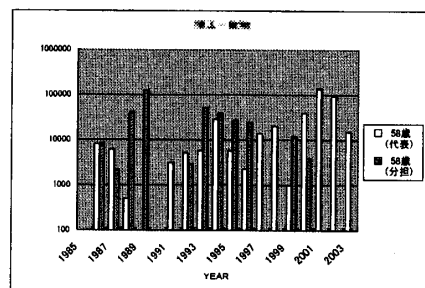
(C) 理工-化学: 理工系の中で化学はやや受給額が多い。科研費の受給割合は84%とやや高くなる。全般的に似た傾向を示し、長年にわたって受給を維持している。積み上げタイプとなる研究者が多いと思われる。また、分担よりも代表の額が2倍近く多いのは分野の特徴であろう。



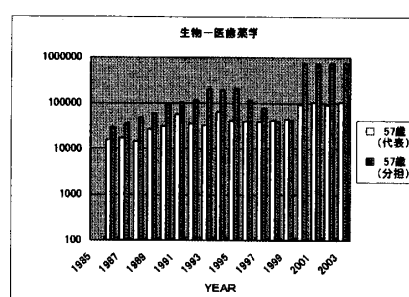
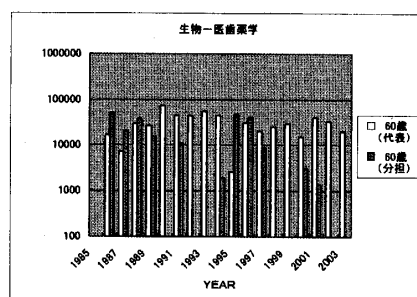
(D) 理工-工学: 理工系の中では、生物系に近いほど配分額が多いが、これは工学系10名中3名が配分額の極めて多い特別推進(COE)の代表者であった事も影響している。全般的に見て高額配分者は多いが、ややばらつきが見られる。このことは科研費受給率が75%と理系ではかなり低い傾向を示したことから伺える。



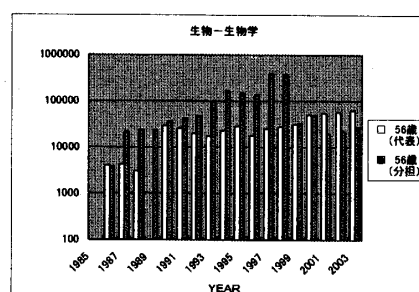
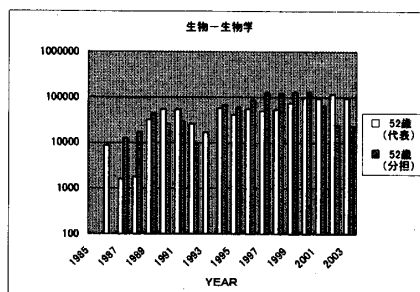
(E) 理工-数物: この分野は科研費依存度が高く、受給率は90%であった。また、化学、工学が代表の額が分担に比べてかなり高かったのと比較すると、ほぼ同額とグループ研究の多さも伺える。ここでの対象者については一部の理論研究者や、海外在住期間が長い研究所を除くと、ほぼ積み上げタイプに分類できる。



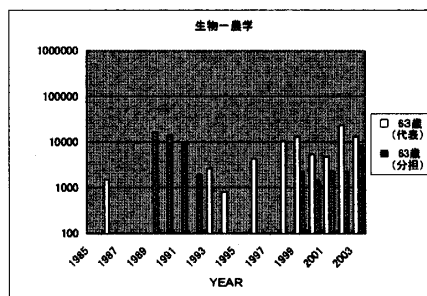
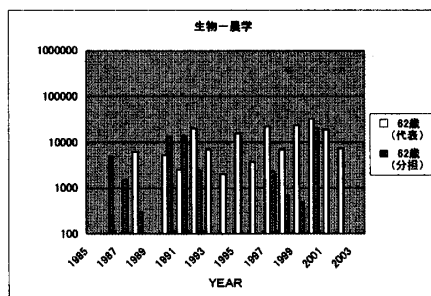
(F) 生物-医歯薬学: この分野は科研費以外の研究費もあるかと思われるが、科研費の依存度は高く96%が受給している。また、代表と分担の比率では、分担額がかなり上回っており、高額なグループ研究の多さが伺える。代表の配分額は生物学に続いて2位である。採択パターンは高額維持タイプが多いが、さらに遡って調査する必要がある。



(G) 生物-生物学: この分野も科研費の依存度は高く94%が受給している。また、代表と分担の比率では、医歯薬学同様分担額がかなり上回っており、高額なグループ研究の多さが伺える。代表、分担共に、受給額は1位である。また、パターンは高額ではあるが、積み重ね傾向が強いところが医歯薬学との違いであろう。

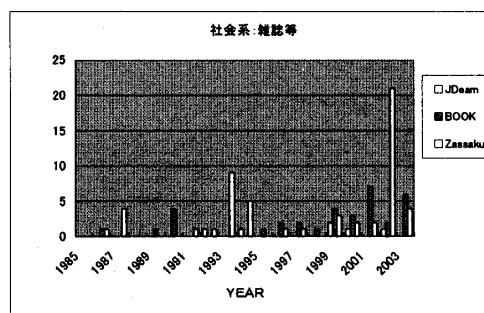
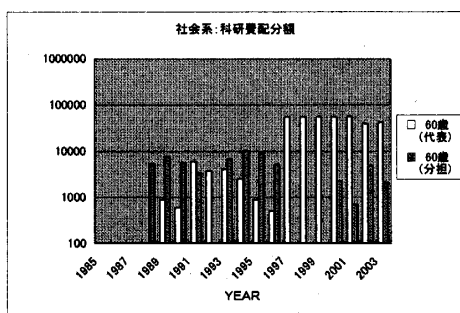


(H) 生物-農学: この分野は高齢の方が多く、65歳以下では15名中2名となった。しかし、科研費への依存度は高く15名全員が受給する100%であった。しかし、受給額自体は理系では最も低い。調査対象が2名でパターン分析は難しいが、積み上げタイプに近い傾向が見られる。



4. 研究成果との関係

今回は、試験的に前回調査した特別推進研究(COE)採択者について、JSTPlus、JMEDPlus、雑誌記事索引等のデータベースから対象者の研究成果を抽出し採択パターンとの比較をおこなった。ISI社のNCR-Jデータベースについても調査中であるが終了していないので影響の少ない社会系の研究者についての一例を示すにとどめる。この一例では、科研費の最終年度と成果発表が相関しているように見える。



4. まとめ

これまでの調査・分析により、各研究者において大型研究の採択に至るまでにいくつかのパターンがあることがわかってきた。また、分野別に見ると分野に特徴的な採択パターンがあるように見える。1965年までの採択課題データベースの遡及入力が終わったので、今後は70歳以下のすべての研究者について調査し、統計的なパターン分析もおこないたい。今後、アンケート調査の結果等、他の方面からの調査結果と統合し、関係の定量化を目指すとともに、研究費と成果の関係についても調査・分析し、より良い研究費配分のあり方について考察していく予定である。

参考文献

- [1]西澤 正己、根岸 正光、柴山 盛生、孫 媛、野村 浩康、光田 好孝、前田 正史、「科学研究費データベースによる採択パターン分析」、情報知識学会誌, Vol.15, No.2, pp.85-88, 2005 (第13回研究報告会講演論文集)
- [2] 2005年4月からKAKENK, KAKENデータベースを統合して科学研究費補助金データベース(KAKEN)としてサービスされている。: <http://seika.nii.ac.jp/>

西澤 正己 国立情報学研究所(〒101-8430 東京都千代田区一ツ橋 2-1-2)
Masaki Nishizawa (nisizawa@nii.ac.jp) National Institute of Informatics

日本の大学と企業の研究リンケージ---NCR-JとCJPデータベースによる比較
Research Linkage between University and Industry in Japan --- Comparison based on NCR-J
and CJP databases

○ 孫 媛, 根岸正光, 西澤正己, 渡辺恵子 (国立情報学研究所)
Yuan SUN, Masamitsu NEGISHI, Masaki NISHIZAWA, Keiko WATANABE
National Institute of Informatics

Nowadays university-industry collaboration is highly promoted in various areas by various ways. In this study we try to evaluate university-industry linkage objectively through analysis of co-authorship of academic publications. To assess the university-industry collaborative activity comprehensively, we choose to use the National Citation Report for Japan (NCR-J) of Thomson Scientific, and Citation Database for Japanese Papers (CJP) of NII as well, to investigate and discuss the situation, trends and characteristics of Japanese university-industry linkages from both international and national perspective.

1. はじめに

筆者らは、日本の学会誌論文を対象として国立情報学研究所が作成した「引用文献索引データベース」(CJP)を用い、主として和文論文における共著関係に注目することで、日本の産学連携の現状分析を試みた(孫ほか, 2005)。本研究は前年の結果を踏まえて、データを更新するとともに、トムソン・サイエンティフィック社(通称 ISI)制作の引用統計データベース National Citation Report for Japan (NCR-J)も併用して、国内の学術誌および国際学術誌への投稿論文のそれぞれについて、産学連携の実態・特徴・動向を分析・比較し、日本の産学連携について国内的および国際的な視点から考察する。

2. データ

本研究では、下記の2種のデータベースを用いた。

2.1 National Citation Report for Japan (NCR-J), 1981-2004

NCR-Jデータベースは、米国トムソン・サイエンティフィック社(通称 ISI)製作の Science Citation Index (SCI), Social Science Citation Index (SSCI), Arts and Humanities Citation Index (A&HCI)という3種類の引用索引データベースに収録された論文から、1981年から2004年に発表された日本人著者を含む論文を抽出して、書誌データと引用数を収録したものである。NCRJ 1981-2004 は1,453,888 件の論文を含むが、レビュー等を除外した原著論文 1,277,823 件を今回の分析対象とする。

2.2 Citation Database for Japanese Papers (CJP), 1995-2004

CJP は、国立情報学研究所が、その前身である学術情報センター時代の1995年から作成している引用文献索引データベースで、日本国内の自然科学分野の学術論文誌・学協会誌論文が収録対象となっている。CJPに収録された文献は、1995年から2004年12月までの1,588論文誌の916,735件に及ぶが、本研究では全文献のうち、①「研究論文・原著・論説」②「レビュー・総説」③「速報・短報」④「ノート」の4文献種(766,652件、全916,735論文の約83.6%)を「研究論文」と

みなし、分析対象とする。それらの文献には、収録雑誌の分類に従って、「理学」「工学」「農学」「医学・歯学・薬学」のいずれの分野分類が付与されている。

われわれは、産官学間の共著論文数を手がかりとして、産官学連携の実態を把握することを試みた。そのために、著者所属機関名によって、全論文を大学、公共機関、企業、外国の4セクタに区分した。ただし、共著論文で著者の所属セクタが複数に及ぶ場合には、関与するすべての所属セクタに重複してカウントした（そのため、各セクタに分類された論文数の合計は、全論文数を上回る）。また、著者の所属機関の名称は原論文での記載のまま収録されるが、表記が多種多様で「ゆれ」が多く含まれているため、セクタへの区分作業をはじめめる前に、各データベースにおける表記のバリエーションを調査し、名寄せ作業を行う必要があった。

3. NCR-Jにおける分析の結果

3.1 セクタ別論文数の経年変化

表1に、1981～2004年の各年度におけるセクタ別論文数などを示す（2004年の数値が前年度より少なくなっているが、データベース作成に一定の作業期間を要することから収録遅れが生じているため、論文数の減少を意味するものではない）。

表1 NCR-Jにおける産公学各セクタ内およびセクタ間の共著論文数

年	実論文数	セクタ別論文数			セクタ内単共著			セクタ間共著			国際共著論文	セクタ間共著率(%)		
		大学	公共	企業	大学のみ	公共のみ	企業のみ	大学+公・産	公共+学・産	企業+学・公		大学	公共	企業
1981	29,128	23,440	4,042	3,952	20,208	2,500	2,852	3,232	1,542	1,100	1,395	13.8	38.1	27.8
1982	29,763	23,991	4,281	4,158	20,247	2,457	2,905	3,744	1,824	1,253	1,656	15.6	42.6	30.1
1983	31,481	25,584	4,504	4,404	21,356	2,482	2,989	4,228	2,022	1,415	1,838	16.5	44.9	32.1
1984	32,970	26,602	4,878	4,804	21,830	2,662	3,183	4,772	2,216	1,621	2,186	17.9	45.4	33.7
1985	35,282	28,404	5,183	5,375	23,116	2,732	3,574	5,288	2,451	1,801	2,432	18.6	47.3	33.5
1986	36,826	29,671	5,649	5,561	23,944	2,959	3,584	5,727	2,690	1,977	2,574	19.3	47.6	35.6
1987	39,411	31,483	6,161	6,297	25,253	3,160	4,029	6,230	3,001	2,268	2,869	19.8	48.7	36.0
1988	41,051	32,912	6,515	6,869	25,732	3,093	4,180	7,180	3,422	2,689	3,302	21.8	52.5	39.1
1989	43,842	35,209	7,128	7,500	27,149	3,281	4,334	8,060	3,847	3,166	3,666	22.9	54.0	42.2
1990	45,595	36,417	7,574	8,152	27,482	3,387	4,627	8,935	4,187	3,525	4,190	24.5	55.3	43.2
1991	49,173	39,006	8,524	9,162	28,821	3,631	5,100	10,185	4,893	4,062	4,876	26.1	57.4	44.3
1992	51,765	41,169	9,187	9,917	29,671	3,628	5,314	11,498	5,559	4,603	5,638	27.9	60.5	46.4
1993	54,465	43,708	9,847	10,468	30,744	3,638	5,323	12,964	6,209	5,145	6,379	29.7	63.1	49.1
1994	58,352	46,841	10,879	11,123	32,354	3,962	5,615	14,487	6,917	5,508	7,313	30.9	63.6	49.5
1995	60,266	48,416	11,837	11,130	32,971	4,198	5,434	15,445	7,639	5,696	8,073	31.9	64.5	51.2
1996	64,350	52,520	12,897	11,361	35,271	4,297	5,304	17,249	8,600	6,057	8,819	32.8	66.7	53.3
1997	66,212	54,419	14,070	11,067	35,600	4,355	4,956	18,819	9,715	6,111	10,010	34.6	69.0	55.2
1998	69,921	57,852	16,305	11,453	35,930	4,520	4,568	21,922	11,785	6,885	11,777	37.9	72.3	60.1
1999	72,466	60,380	17,868	11,292	36,683	4,750	4,172	23,697	13,118	7,120	12,697	39.2	73.4	63.1
2000	73,908	61,988	18,977	10,845	36,807	4,970	3,726	25,181	14,007	7,119	13,665	40.6	73.8	65.6
2001	73,305	61,633	19,698	10,414	35,734	4,920	3,372	25,899	14,778	7,042	14,286	42.0	75.0	67.6
2002	74,359	62,773	20,552	10,501	35,337	5,090	3,179	27,436	15,462	7,322	15,039	43.7	75.2	69.7
2003	76,340	64,755	21,832	10,545	35,448	4,967	2,975	29,307	16,865	7,570	16,155	45.3	77.2	71.8
2004	66,799	56,901	19,791	9,035	30,247	4,332	2,313	26,654	15,459	6,722	14,810	46.8	78.1	74.4

「セクタ内単共著」は、その論文の著者すべて（単著を含む）が同一のセクタに属するものを指す。「セクタ間共著」は異なるセクタに所属する著者を含むもので、たとえば「大学+公・企」は、大学研究者と公共または/および企業研究者との共著論文のことである。これらは該当各セクタに重複して計上されている。各セクタの論文に占めるセクタ間共著論文の比率は「セクタ間共著率」に示されている。「国際共著論文」は海外機関所属の著者を含む論文である。

この20年余を概観すると、論文数は着実に増加しており、国際的視点からわが国の研究水準の向上が示される。ただし、企業関連の論文数は1998年を頂点として漸減傾向にあるのが目を引く。この間国際共著論文が激増している点も注目され、わが国の研究活動の国際化の一端が示されて

いるものと考えられる。また、セクタ間共著率の変化に着目するならば、大学で14%から47%、公共で38%から78%、企業で28%から74%というように、セクタ間共著の比率が大きく増えていることがわかる。

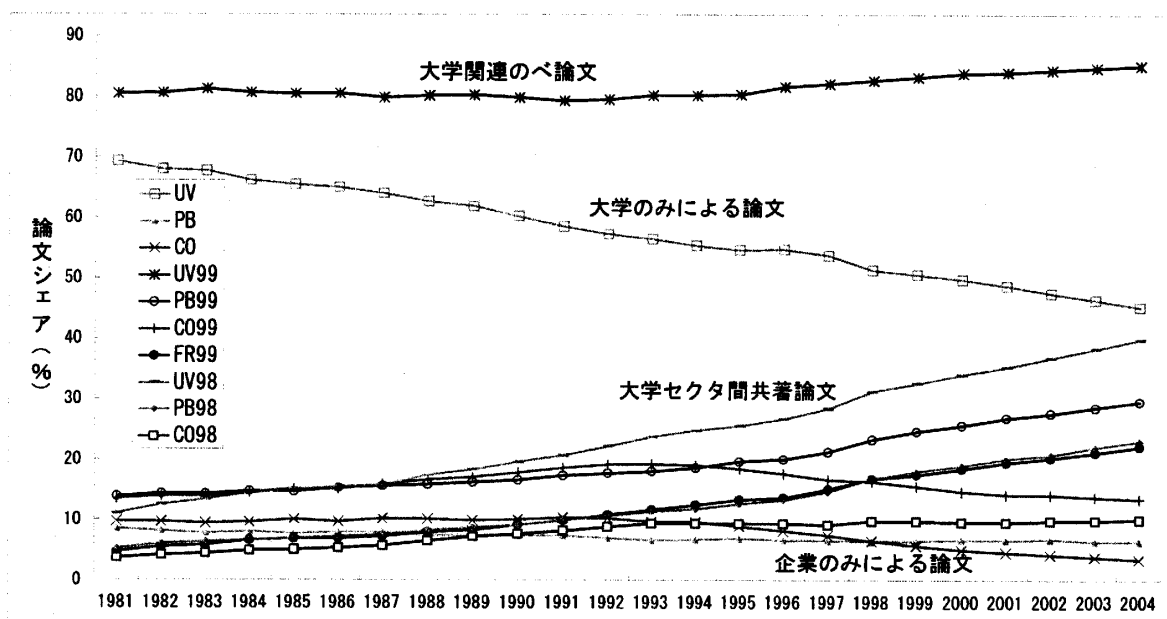


図1 NCR-Jにおけるセクタ別論文シェアの推移

図1は、大学、公共、企業のセクタ別に、セクタ内・間共著論文、セクタ内論文、セクタ間論文のシェアの推移を示したものである (UV, PB, COはそれぞれ大学等、公共、企業の各セクタ内単共著、99が付されたものはセクタ内・間共著合計、98が付されたものはセクタ間共著のみである)。大学について見ると、大学内・間共著論文 (UV99) がここ10年ほどで一層比重を増し85%程度を占めていること、しかし大学セクタ内論文 (UV) にかぎると、シェアは70%から45%へと著しく下落していることがわかる。そのかわりとして増えているのが、企業や公共機関との共著シェア (UV98) である。一体に、単独セクタ内論文のシェアが低下し、セクタ間共著が増加する傾向にある。

3.2 共著相手構成比の経年変化

大学・公共・企業のそれぞれについて、当該セクタ関連共著論文の総数を100として、共著相手の構成比を示したのが表2である。

表2 NCR-Jにおける産公学各セクタ別関連論文における共著相手の比率 (%)

年	大学関連論文				公共関連論文				企業関連論文			
	大学のみ	公共との共著	企業との共著	海外との共著	公共のみ	大学との共著	企業との共著	海外との共著	企業のみ	大学との共著	公共との共著	海外との共著
1981	86.2	5.5	3.8	5.1	61.9	31.9	5.0	4.1	72.2	22.3	5.1	2.7
1985	81.4	7.2	5.0	7.3	52.7	39.4	6.0	5.8	66.5	26.5	5.8	4.2
1990	75.5	9.2	7.7	9.4	44.7	44.2	8.9	9.1	56.8	34.2	8.3	5.8
1995	68.1	12.6	9.1	13.2	35.5	51.6	10.3	13.2	48.8	39.6	11.0	9.4
2000	59.4	18.6	9.0	17.9	26.2	60.9	9.9	17.4	34.4	51.6	17.4	13.0
2003	54.7	21.6	9.3	20.4	22.8	64.0	9.8	19.4	28.2	57.1	20.2	14.8
2004	53.2	22.7	9.3	21.4	21.9	65.2	9.9	19.9	25.6	58.6	21.7	16.0

大学側から見ると、セクタ内 (大学のみ) の論文の比率は1981年の86%から減り続け、近年では50%台となっている。目立って増えているのは、公共セクタおよび海外との共著論文の割合で、

最近ではいずれも20%を超えている。企業との共著は増えてはいるものの、9%台に止まっている。公共側から見ると、セクタ内（公共のみ）の比率は減っており、大学および海外との共著の割合が増えている。とくに大学との共著比率は60%を超えている。また、企業側から見た場合も、セクタ内（企業のみ）の論文は1981年の70%超から約25%まで減り、大学との共著論文は60%に迫る勢いである。この20数年の間にセクタをまたがった共著論文の割合が増していることは明らかであり、とくに産官にとって研究の協力相手として大学の地位が高まっていることがわかる。

4. CJP における分析の結果

前節に示した NCR-J にもとづく調査分析は、国際学術コミュニティにおける日本の投稿論文の状況、すなわち国際視点からみた日本の学術研究および産官学連携の状況を示すものと考えられる。それに対して、CJP データベースを用いた分析は、日本の国内学術雑誌論文における産官学研究協力に関する分析となる。

4.1 セクタ別の論文比率

「理学」「工学」「農学」「医学・歯学・薬学」の分野別に、各セクタに分類された論文数をまとめたものが表3、各分野の実論文数を100としたときの比率を示したものが表4である。

表3：CJPにおける分野別にみたセクタ別論文数

分野	論文数						
	大学	企業	公共	外国	その他	未調査	実論文数
理学	68,023	5,304	13,482	6,793	1,304	28,465	98,585
工学	196,065	93,060	38,449	14,492	1,522	74,724	328,728
農学	43,278	5,013	16,373	2,233	446	21,567	68,431
医学・歯学・薬学	183,667	10,328	71,186	10,791	3,758	68,514	270,908
総計	491,033	113,705	139,490	34,309	7,030	193,270	766,652

表4：CJPにおける分野別にみたセクタ別論文比率

分野	実論文を100としたときの比率 (%)						
	大学	企業	公共	外国	その他	未調査	総計
理学	69.0	5.4	13.7	6.9	1.3	28.9	125.1
工学	59.6	28.3	11.7	4.4	0.5	22.7	127.3
農学	63.2	7.3	23.9	3.3	0.7	31.5	129.9
医学・歯学・薬学	67.8	3.8	26.3	4.0	1.4	25.3	128.5
総計	64.0	14.8	18.2	4.5	0.9	25.2	127.7

大学セクタの論文総数は491,033件で、実論文数(766,652件)の64%を占めている。分野別に見ても、実論文数に占める比率は大学セクタが高く、最小の工学でも59.6%、最大の理学では69%を占めている。企業セクタ論文の比率は、理学、農学、医学・歯学・薬学の3分野では10%に満たないが、工学分野では28.3%で、他分野よりもかなり高い値を示している。公共セクタの比率は、理学あるいは工学と比べて、農学と医学・歯学・薬学で高くなっている。

セクタ別の論文数を年度別にまとめたものが表5である。NCR-Jの結果と比べたとき、この10年間における各セクタ、セクタ内・間に大きな変化がないことが見て取れよう。実論文数に占める割合は、大学は約45%、公共と企業はそれぞれおよそ8%台を占めている。

表5 CJPにおける産公学各セクタ内およびセクタ間の共著論文数

年	実論文数	セクタ論文のべ計			セクタ内論文			セクタ間論文			国際共著論文
		大学	公共	企業	大学のみ	公共のみ	企業のみ	大学+公・産	公共+学・産	企業+学・公	
1995	69,264	31,750	6,131	6,689	29,522	8,009	7,669	6,863	3,989	3,474	443
1996	68,928	31,334	6,106	6,429	29,344	8,070	7,326	6,918	4,122	3,344	466
1997	72,349	32,135	6,519	6,998	30,739	8,542	8,100	7,295	4,217	3,741	545
1998	73,624	32,126	6,904	6,645	30,931	9,207	7,765	7,558	4,550	3,647	602
1999	80,432	35,548	7,355	6,758	34,772	10,123	8,022	8,436	5,172	3,845	638
2000	85,276	37,436	7,624	5,980	35,806	10,209	7,347	8,545	5,178	3,803	705
2001	92,236	40,031	8,321	6,775	38,158	11,000	8,490	9,115	5,521	4,291	744
2002	86,761	38,030	8,131	6,125	35,999	10,685	7,233	8,551	5,414	3,646	757
2003	78,564	34,564	6,919	5,906	33,215	9,374	7,064	7,635	4,762	3,369	671
2004	59,218	26,005	5,092	4,796	25,049	6,520	5,598	5,425	3,236	2,498	513

4.2 セクタ別論文比率の経年変化

図2と図3は大学関連と企業関連の論文比率の変化を示している。図2に示されるように、大学関連の論文では、大学のみ（単著と共著を含む）による論文が全体の80%を超えており、大学と企業の共著論文は工学分野での20%弱を除くと10%未満である。一方、企業関連の論文を見ると（図3）、いずれの分野においても大学との共著論文の比率が高く、企業単独の比率が高い工学分野でも40%前後、他の3分野では40%を超えて、さらに増加の傾向にある。とくに理学で大学との共著論文の割合が企業のみ（単著と共著を含む）の論文比率を常に上回り、その差はますます拡大している。

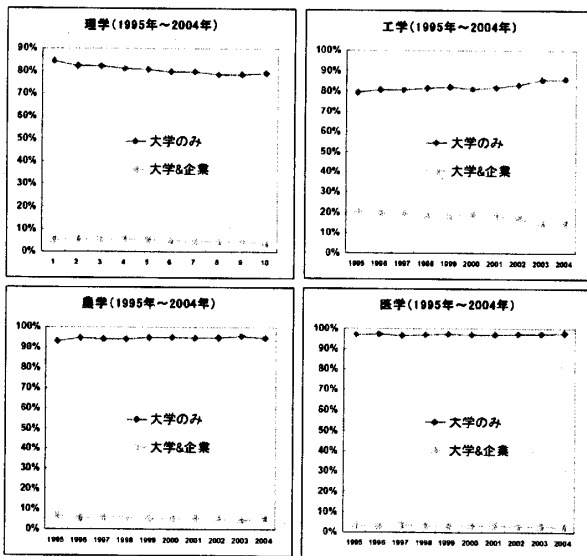


図2 大学関連論文比率の推移

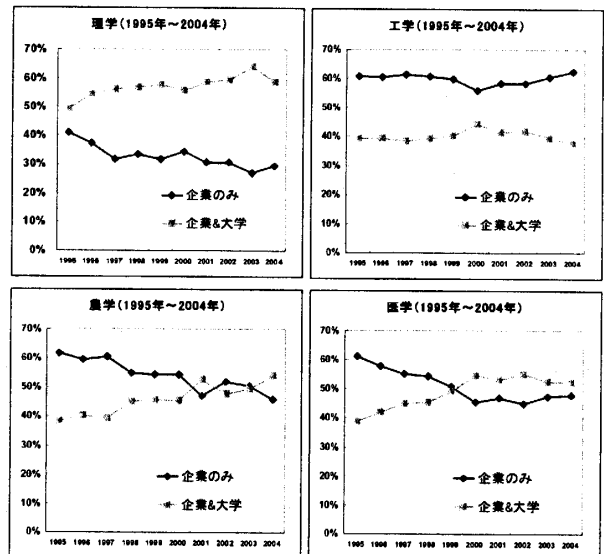


図3 企業関連論文比率の推移

NCR-Jの結果と同様に、企業にとって大学との共同研究は今や研究開発における最大・不可欠の要素ということを示している。

5. おわりに

NCR-J, CJPいずれにおいても、セクタをまたがる共著の増加が示された。大学側から見てもそうだが、ことに公共、企業の論文に占める大学との共著の割合が大きく増えており、大学が産官学の共同の中心的役割を担いつつあるように見える。海外の学術誌への投稿論文に関するNCR-Jにおいてその傾向は顕著であった。この20年間の産官学の連携の進展の背景には、それを進める

ための制度の整備があったと考えられる。そこで最後に、産学連携に関わる制度の変遷についてまとめておきたい。

産学連携に関わる制度の変遷を見るときに、2つの画期があることが指摘できよう。一つは科学技術基本計画が制定された1996年、もう一つは国立大学が法人化された2004年である。以下、この時代区分を基にしてそれぞれの時代における制度の変遷を辿る。なおここでは、勤務条件や会計ルールが国の機関として法定されており、産学連携促進のために法令改正が必要であった国立大学を対象とした制度の変遷を取り上げることにより、それぞれの時代における特徴にも付言したい。

(1) 1995年以前

1983年に、国立大学等において企業等から研究経費等を受け入れて共同して研究を行うための制度としての共同研究制度が創設され、その後、旧来から可能であった奨学寄附金や受託研究費の制度も含めて徐々にその取扱いの弾力化が行われた。また、1987年には、企業等からの寄付に基づいて講座や研究部門を設けることができる寄附講座・寄付研究部門制度が創設された。この期間は主に、企業の意向を反映できる形態での企業からの経費の受入れを行う制度を整備するとともにその経費の運用を改善することによって産学連携の促進を図ることに力点があったといえよう。

(2) 1996年から2003年まで

1996年に制定された科学技術基本計画においては、産学連携に関し、人的交流の促進と研究成果の民間における活用の促進が謳われた。このため、この期間はこれらの点に関する制度改正が主に行われた。人的交流の促進の観点からは、1997年に大学教員の任期制が導入されるとともに、国立大学教員等が技術コンサルティング兼業を行うことが可能となり、さらに2000年にはTLOや自らの研究成果を活用して事業を行う企業などの役員に就任することも可能となった。研究成果の民間における活用の促進の観点からは、共同研究や受託研究において相手先企業へ特許権の優先実施権を付与することが可能となり、1998年に承認TLO制度が創設されるとともに、1999年にはいわゆる日本版バイドール法が制定された。

(3) 2004年以降

2004年に国立大学が法人化されたことにより、勤務条件や会計ルールに関する国の規定の適用がなくなり、各大学でルールを定めることが可能になった。また、法人格の取得により特許等を各大学に帰属させることも可能になった。これらの制度の変更により、今後、各大学が産学連携のために取りうる選択肢が一層多様になり、大学ごとに制度が異なってくることが予想される。

参考文献

- 孫媛，西澤正己，根岸正光「日本の論文の共著関係からみた産学連携の現状分析」，情報知識学会誌，vol.15, No.2, pp.89-92, 2005
- 根岸正光「大学の研究における産学連携を統計的に分析する」有本章，北垣郁雄編（広島大学高等教育研究開発センター）「21世紀の大学力」，ミネルヴァ書房，2006
- 伊藤学司「わが国の産学連携の現状と課題」『応用物理』第75巻第1号，2006
- 国立大学等外部資金取扱事務研究会編著「大学と産業界との研究協力事務必携〈第四次改訂版〉」ぎょうせい，2001

ISIデータベースにおける発表論文数、被引用回数に基づく
「大学ランキング」作成の方法と問題点

Formulating methods of university rankings by publication and citation statistics
in ISI's citation databases, and some related problems

根岸正光（国立情報学研究所）

Masamitsu NEGISHI, NII: National Institute of Informatics

Since 2001 the author has been compiling university rankings for "The University Rankings," an annual publication by the Asahi Shimbun Company, based on statistical surveys on the ISI citation databases of Thomson Scientific. In this report, the historical developments in these 6 years are reviewed, and a new method of calculating number of papers and citations introduced for the 2006 edition is described with some related problems which are involved in these bibliometric statistics.

1. はじめに

筆者は、2001年以來、トムソン・サイエンティフィック社（通称ISI）制作の引用統計データベースNational Citation Report (NCR)を用いて、わが国の大学別の発表論文数およびそれらへの被引用回数に基づく大学ランキングを、朝日新聞社刊『大学ランキング』誌の各年版に公表してきた《表1》<1>。2006年4月発刊の最新版では、大学間共著論文を大学別に案分して集計し（案分論文数）、また、引用度指数の算定においては、論文当たり引用数の大学間格差をより正確に反映できるとみられる偏差値（引用度偏差値）を導入したランキングの生成を試みた<2>。本報告では、これまでのランキング作成手法の改変を回顧しつつ今回導入した新手法を解説し、その効果あるいは副作用を分析して、この種の計量書誌学的番付に関わる問題点を論じる。

2. ランキング作成方式の改訂経過

表1 『大学ランキング』誌での大学別論文数・引用度調査結果掲載状況

『大学ランキング』誌版次	2002	2003	2004	2005	2006	2007
発行年月	2001.5	2002.5	2003.5	2004.5	2005.4	2006.4
標榜掲載大学数	682	699	710	717	726	?
調査対象年次	81-97	90-99	92-01	93-02	03-02	00-04
調査対象年数	17	10	10	10	10	5
調査対象論文数	740千	483千	549千	573千	573千	310千
配列順序	論文数	総合引用度 指数	総合引用度 指数	総合引用度 指数	総合引用度 指数	案分加重引 用度偏差値
引用度計算法	論文当引用数	同上	同上	同上	同上	同上
掲載頁数	14	8	8	8	4	8
総合表掲載大学数	140	140	139	140	69	70
掲載分野数	24	20	20	20	10	20
NCR統計年次	81-97.6	81-99	82-01	81-02	81-02	81-04
NCR収録総論文数	85万	106万	120万	133万	133万	145万

『大学ランキング』誌は、1994年4月刊行の「95年版」を創刊号として、毎年春に刊行されて現在に至っている。筆者は同誌編集者から依頼を受けてその2002年版から、前記NCRデータベースによる調査結果を公表しており、2006年刊行の2007年版で6回目ということに

なる《表1》。なお、2005年版と2006年版は同一のNCRデータによる同一の結果である。

ここでNCRとは、ISIが作成している引用索引データベース、SCI(オンラインサービス名はWeb of Science)、SSCI、A&HCIに収録された各論文に対する、他の論文からの現時点までの引用回数を集計して入力した引用統計データベースで、名称のとおり、論文著者の所属機関の所在国に応じて国別に作成される。われわれが用いているのはその日本版で、より正確には“NCR for Japan”である。ここには、各論文について、著者、所属機関、標題、掲載誌、巻号頁、分野分類(100種。筆者の調査ではこれを20分類に集約して統計している)、引用回数等が収録されている。

大学ランキングのための大学別集計は、著者所属機関欄の記述に即して行うのであるが、これは原論文での記載をそのまま略記入力したものであり、記法のゆれが大きい。端的な例は、東京大学の正規英文名“Univ of Tokyo”に対する“Tokyo Univ”等々であって、これらの表記異同についてデータベース内での実態を調査しつつ、大学別の名寄せ、集計を行っており、これは相当の工数を要するものになっている。また、昨今は大学の統合再編、名称変更などが盛んで、これについても同様に調査の上集計している。

『大学ランキング』誌における、この調査結果の最初の掲載は2002年版で、これはNCRに対する調査結果として、筆者らがわが国初の大学別・分野別の論文数、引用数を公表したことを受けたものである<3>。ここではまず配列順を、東大を先頭として基本的に大学の規模順となり、そのいみで「穏便な」論文数順とし、これに引用度として単純な論文当たり引用数を付加した体裁とした。入試偏差値ランキングは昔からあったが、論文数、引用度など、大学の研究評価に関わる番付は、この当時、特に大学関係者から忌避されるような雰囲気が強かったからである。しかし、こうした状況は2001年6月の文部省の「トップ30大学構想」により一変し、大学間の競争が露骨に語られるようになり、この傾向は2004年の国立大学の法人化以降さらに顕著になっている。

こうしたことから、翌2003年版では配列順を論文品質に関わる引用度順にすることにし、また統計期間も最近の研究状況を反映させるべく、10年間に短縮した。引用度順にすれば、読者の関心はどうしてもこれに集中するので、その算定法には工夫を要する。そこで、後述のとおり分野間における平均引用度の格差を折込み、同時に各大学の研究分野構成を反映した「総合引用度指数」を考案し、これによる番付を作成、公表した。2006年版までこの方式を維持したが、最新の2007年版では、さらに「案分論文数」と引用度の偏差値を導入した「案分加重引用度偏差値」を考案し、これによる順位表を公表することにした。

3. 総合引用度指数—2003～2006年版

引用数を論文数で割った単純な引用度には分野間で大差がある。例えば、分子生物学・遺伝学では分野内の平均引用度は15.7に達するが、工学分野では2.8程度である(1990-1999年)。これは論文のスタイルに依存するもので、論文末尾に掲げる引用論文リストの長さが、生物系では50件以上でも普通なのに対し、工学系では10数件程度といった状況に起因する。この引用度を単純に受け取ると、生物系の研究はいかにも進んでいるが、工学系の研究水準は低位に止まっているという解釈になり、今後わが国はあげて生物系研究に重点をおいた科学技術政策を展開すべしといった主張につながるし、実際、このような「勇ましい」論調も散見される。

そこで2003年版では、20の分野別に大学別引用度および分野内全体の平均引用度を計算し、前者を後者で割った指数、「引用度指数」を算出した。これによれば、指数100の大学はわが国での平均的水準の論文を生産していることになり、100以上の大学は平均以上の論文品質を保持していることが分かる。

つぎに、各大学の分野別引用度指数を分野別論文数で加重平均して、「総合引用度指数」とし、これによる番付を「総合ランキング」として掲げた。これにより、総合大学、単科大学を通じて、各大学のわが国における相対的位置な関係が表現されたものと思われる。

4. 案分加重総合引用度偏差値－2007年版での計算法改訂

引用度偏差値－ところで、総合引用度指数によると総体的に医学系単科大学の数値が高く出がちであるという傾向が見られる。これには、分野間での引用度分布の形状の差異が影響しているのではないかと考えられた。すなわち、医学系では、一部の単科大学に引用度の突出したものがあり、これは分野内比較においては問題を生じないが、論文数による加重平均で総合化する際に、それら突出した数値が医科単科大学に「有利」に、また医学部を含めた多学部を抱える総合大学には「不利」に働くのではないかと。

表2 『大学ランキング2007年版』掲載の総合ランキング（上位大学のみ）

引用順位	大学名	論文数	共著案分後論文数	総合引用度偏差値	引用順位	大学名	論文数	共著案分後論文数	総合引用度偏差値
1	基礎生物学研究所	597	269.0	152.7	16	東北大	18830	12857.6	112.3
2	国立遺伝学研究所	558	273.7	136.9	17	名古屋大	12409	7934.0	112.2
3	生理学研究所	579	281.0	132.0	18	宇宙航空研開機 構宇宙科学本部	1113	610.5	111.9
4	奈良先端大	1600	964.3	125.4	19	中央大	584	349.2	111.9
5	高エネ研究機構	2648	1217.8	121.8	20	旭川医大	836	552.9	110.9
6	都立大（首都大）	2708	1620.2	120.3	21	国立天文台	1281	733.6	110.5
7	東京大	32213	19904.0	118.1	22	名城大	946	596.8	110.4
8	総合研究大学院大	1235	541.2	116.4	23	横浜市立大	2104	1295.3	110.2
9	京都大	22766	15347.2	116.3	24	兵庫県立大	1684	1076.6	109.8
10	京都薬大	802	577.0	115.8	25	東京工業大	11364	7973.8	109.7
11	分子科学研究所	1506	772.4	115.5	26	東京医歯大	3503	2122.0	109.5
12	大阪大	19918	13299.4	114.8	27	東京薬大	1004	693.6	109.5
13	順天堂大	1881	1186.0	114.1	28	筑波大	8247	5074.5	109.4
14	青山学院大	533	299.2	114.0	29	北陸先端大	1054	771.8	109.0
15	関西学院大	598	422.9	113.2	30	久留米大	1449	935.8	108.7

2000-2004年。ISI NCR for Japan (1981-2004)に対する根岸の調査結果。

この点を「是正」するために、2007年版では引用度偏差値を算出してみた。すなわち、各分野について、大学別の単純引用度の偏差値を計算して分布を標準化した後の値を算出した。この場合、従来の指数との感覚的親和性を考慮し、種々試算の結果、平均100、標準偏差30とする偏差値に換算した（一般のいわゆる偏差値は、平均50、標準偏差10）。これにより分野間における引用度分布の形状の相違は一応吸収される。

案分論文数－自然科学系では共著論文が多いが、研究評価におけるその扱い方もよく問題にされることである。この場合最も「正確」な計算法は、著者数に応じて論文数を配分する方法で、例えばA大学の著者3名、B大学1名、C企業1名の共著論文は、A0.6件、B、C各0.2件のように案分することになる。しかし、NCRでは原データの構成上この計算は不可能なので（著者と所属機関を個別に対応させた記述でなく、所属機関の一覧を別途掲げる）、ここではA、B、C各々に0.33件と割り振る計算を試みた。

案分加重総合引用度偏差値－そして、分野別指数の加重平均による総合指数を出す際の

重みとして、この案分論文数を用いることにした。これを「(案分加重) 総合引用度偏差値」と称することにする《表2》。

統計期間—いつからいつまでの論文を調査対象とするべきかについても、従来から議論のあるところで、最新の研究活動を反映させるという観点からは、最近数年間の論文に限定するべしということになる。もっとも引用が蓄積され、傾向が安定するにはある程度の年数を要するので、例えば昨年発行の論文のみに限定するなど是非合理である。これまで統計期間を10年間としてきたが、こうした情勢を受けて、今回はこれを5年間に短縮した。もっとも、論文の「寿命」は分野間での差異が大きく、やはり10年程度が適当とも考えられるところで、別途10年統計も作成してみたい。

5. 引用統計の問題点

前記のとおり、この種の引用統計では、論文引用における分野間格差の処理方法、共著論文の案分処理の問題、論文寿命と関連する統計期間の適切性の問題がある。しかしそれにも増して大きな問題は、論文収録範囲の問題であろう。

この引用統計はISIのデータベースによるもので、これは全世界から有力な学術誌を「厳選」し(自然科学系で5000誌程度)、その掲載論文のデータを収録したもので、基本的に英文誌となっている。従って、わが国の和文学会誌を含めて、ISI非収録の学術雑誌の方が圧倒的多数を占めるのは当然である。この際人文社会系はさておくとして、わが国の場合、特に工学系では和文学会誌への評価には高いものがあるが、これらは統計の埒外になっている。このことから、工学系ではISIデータベースによる論文統計には反発も強い。一方、海外英文誌への投稿が旺盛な生物医学系では、ISIの統計で十分であり、また雑誌が国際的に「権威ある」ものだけに厳選されているからこそ有意義であるとの見解もあって、議論がかみあわない状況が続いている。

自らに都合な統計数値はせいぜい「活用」し、不利な数字は批判するというのは人間的に自然な行動であろうが、この際科学者として、統計の背景、性格、特性に関して一定の理解の上に立っての議論が望まれるところであろう。

6. 案分論文数・引用度偏差値導入の効果分析

表3 共著論文の機関別案分による論文数の目減り効果

目減り率順位	大学名	論文数	共著案分後論文数	総合引用度偏差値	案分目減り率	目減り率順位	大学名	論文数	共著案分後論文数	総合引用度偏差値	案分目減り率
1	総合研究大学院大	1235	541.2	116.4	0.44	11	青山学院大	533	299.2	114.0	0.56
2	基礎生物学研究所	597	269.0	152.7	0.45	12	藤田保健衛生大	929	528.2	105.7	0.57
3	高エネ研究機構	2648	1217.8	121.8	0.46	32	東京大	32213	19904.0	118.1	0.62
4	生理学研究所	579	281.0	132.0	0.49	92	京都大	22766	15347.2	116.3	0.67
5	国立遺伝学研究所	558	273.7	136.9	0.49	127	名古屋工大	2050	1490.6	97.2	0.73
6	奈良女子大	791	403.7	103.3	0.51	128	豊橋技科大	1399	1022.5	107.1	0.73
7	分子科学研究所	1506	772.4	115.5	0.51	129	北陸先端大	1054	771.8	109.0	0.73
8	愛知医大	664	352.9	100.3	0.53	130	防衛大学校	637	467.9	101.9	0.73
9	宇宙航空研開機構	1113	610.5	111.9	0.55	131	長岡技科大	1104	816.0	98.0	0.74
10	川崎医大	733	404.0	99.9	0.55	132	大分大	1282	971.7	99.6	0.76

目減り率—前記のように論文数を共著大学別に案分計算した結果と、通常の論文数(共著各機関とも1件ずつとして重複計上される)との比率、「案分目減り率」を、論文数上位の132大学について計算すると《表3》のようになる。平均的目減り率は0.64となり、その

上位には総合研究大学院大学と基礎生物学研究所等、その「基盤機関」がならぶ。総研大は大学共同利用機関(基盤機関)において大学院教育を実施するために設立されたもので、形式上研究者は併任で、論文には所属機関として両者を併記することになるから、こうした結果が出て自然である。これをあえて回避するには総研大・基盤機関間の共著は案分に算入しないとす特別な処理が必要である。

これ以外の目減り上位大学は奈良女子大、愛知医大、川崎医大、青山学院大、藤田保健衛生大、埼玉医大等であるが、医大では、その大学付属病院以外の各地の病院との共著も多いという結果ではないか。また奈良女子大等では、大規模大学との共著が多いのでなかろうか。一方目減りの少ない、独自の研究発表が多い大学は、大分大、長岡技科大、防衛大、北陸先端大、豊橋技科大、名古屋工大などである。防衛大は当然として、その他比較的小規模な工科大では、学内で独自の研究が行われていることをいみし、この点は興味深いのでさらに調査してみたい。ともあれ、大学単位の共著・単著の率については、その共著相手に応じて、例えば独立独歩型、大規模大学連携型、地方連携型等々の類型化が可能と思われ、各大学の研究推進戦略と合わせて今後検討してゆきたい。

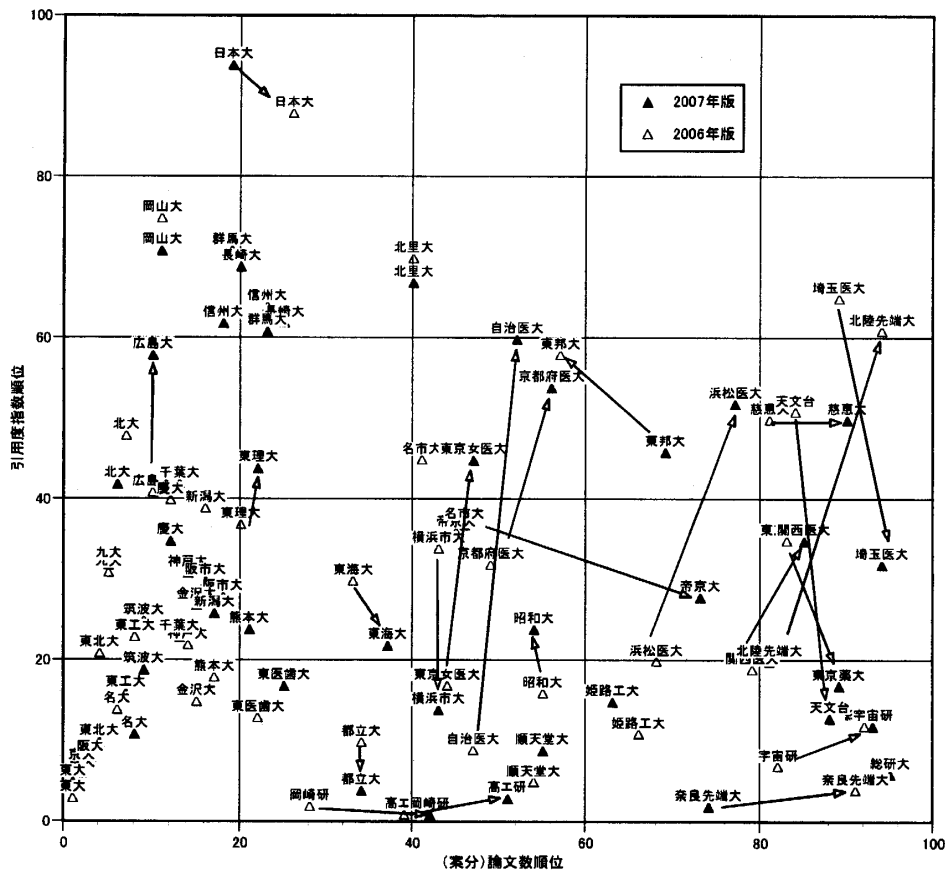


図1 案分論文数、引用度偏差値導入後の順位変動

2006年版からの順位変動-2007年版で導入した引用度偏差値、案分論文数、そして統計期間の5年間への短縮について、その作用、効果を概観するために、順位に着目して2006年版との間でのその変動をみってみる(同一のNCRデータによる正確な比較分析は今後実施したい)。ここでは、2006年版の論文数順位、総合引用度指数順位、2007年版の案分論文数順位、案分加重総合引用度偏差値順位の4変数により、両年の論文数、引用度と案分目減り率において上位にくる特徴的な大学を抽出して散布図を描いてみると《図1》のようになる。

一見して、論文数の多い大規模大学では大きな変位はみられないが、論文数の少ない大学では引用度順位が大幅に低下しているものがある。特に医科大ではその傾向がみられ(埼玉医大は逆)、これは引用度偏差値を導入した効果と考えられる。また既述のとおり旧岡崎研究機構、高エ研や宇宙研は総研大との関係で案分による目減りが大きく論文数順位を下げている。これら変動の距離と方位を中心点から放射状に配置して《図2》のような極座標グラフを描くこともできる。ここで、右上方向は論文数、引用度ともに順位の低落、左下方向は順位の上昇を表すことになる。これにより注目すべき大学を抽出して、変動の要因をさら詳細に分析することができるであろう。

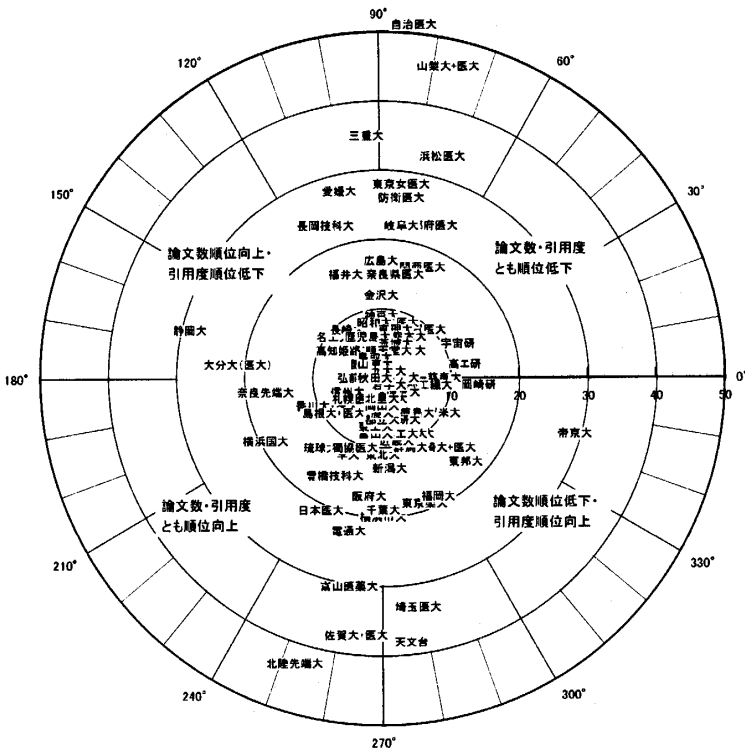


図2 極座標グラフによる順位変動表示

7. おわりに

今後ともこのような試行と分析を重ねることにより、より実態に即したランキングにしてゆくことができると思われる。それにしても、この種の番付は、その作成方式に応じて各々特性、あるいはバイアスをもつものであるから、利用者における誤解、曲解を排除するよう、その特性等についてよく説明してゆくことは作成者の責任といえる。一方、利用者の方でも自分に好都合な数値を無闇に振り回すのではなく、統計の作成法と性格に対して理解を深める努力も肝要である。研究評価に

基づく重点的予算配分がいわれる今日的状況においては、なおさらしかりであろう。

この大学ランキングの作成方法について、名古屋大学大学院理学研究科物理学教室の岡本祐幸教授には、分子科学研究所職の当時から有益な助言を頂いてきており、既述、偏差値導入の件もその一端である。記してここに感謝したい。

〈1〉 根岸正光「日本人研究者の論文85万件から分野別の研究勢力を探る」、「大学ランキング2002年版」(朝日新聞社)、p. 64-77 (2001. 5); 「論文の引用度から分野別の研究勢力を測る」、「同2003年版」、p. 134-141 (2002. 5); 「ISI・論文引用度指数ランキング: 引用が好対照な小柴昌俊さんと田中耕一さん」、「同2004年版」、p. 156-163 (2003. 5); 「ISI・論文引用度指数ランキング: 引用度の高い論文から大学の得意分野がわかる」、「同2005年版」、p. 186-193 (2004. 4); 「ISI論文引用度指数ランキング: 57万3000件の集計から日本の研究者の実力を知る」、「同2006年版」、p. 204-207 (2005. 4)

〈2〉 根岸正光「ISI論文引用度指数ランキング」、「大学ランキング2007年版」(朝日新聞社)、8p. (2006. 4刊行予定)

〈3〉 根岸正光、山崎茂明編著「研究評価 — 研究者・研究機関・大学におけるガイドライン」、丸善、2001. 5、204p. ISBN4-621-04890-2

最適解の付与に基づく重み付けの自動変更

Assigning kansei keywords automatically based on the optimal keywords which experts assigned

原田隆史(慶應義塾大学文学部)

HARADA, Takashi, School of Library and Information Science, Keio University

In recent years, some novel retrieval systems which use kansei keywords have been developed. We have developed the retrieval system that 11432 children's books can be searched for by the kansei keywords picked up from online book reviews. The system was evaluated by school librarians who found that the system was useful for the education of children. In order to improve the system to the level of practical use, to add kansei keywords is necessary, few experts, however, can deal with kansei keywords. The purpose of this research is to examine the methods of adding kansei keywords automatically based on the keywords which experts added. In case the words are in common with the words in the book review the experts added the kansei keywords, the same kansei keywords are added the book.

1. 日本の状況に対応した感性語による図書検索システム

学校図書館において児童生徒が読みたい図書を求める際には、書名や著者名といった項目からだけではなく、「悲しい本が読みたい」「おもしろい本はないですか」などのように読後感からの検索要求を示すことも多いとされる¹⁾。このような図書を探す際の多様な読書ニーズに対応するためには、個々の本が待つ雰囲気や印象、読後感などを覚えておき、読書指導の場面で必要となった時に対応するようにすることが望ましい。

図書から感じる感情や雰囲気など幅広い図書の評価から検索することができるシステムとしては、英国のwhichbook.netがある²⁾。しかし、whichbook.netは英語の図書を対象としており、当然のことながら日本語の図書には対応していない。また、総合的な学習の時間など日本独自のシステムへの対応なども考慮する必要があるため、日本の図書を対象としてwhichbook.netをそのまま使用することはできない。そこで著者らは、whichbookをモデルに、図書の雰囲気・図書から受ける印象を表す感性語をもとにした図書検索が可能な日本独自のシステム「日本語版whichbook(仮称)」を作成を行っている³⁾。作成しているシステムは、英国で作成されたwhichbook.netシステムを参考に新しく開発したもので、英国版のプログラムの提供を受けたものではない。英国のwhichbook.netとの主たる相違点は以下の通りである。

- 1) 日本で刊行された図書を対象としている。現時点では、オンライン書店であるアマゾンコム¹⁾の分類で児童書・ヤングアダルト図書に分類されていた図書およびヤングアダルト図書総目録2004²⁾に収録された図書合計22,765冊のうち、オンライン書店ア

マゾンコム¹⁾またはビーケーワン⁶⁾に書評が掲載されていた11,432冊を対象とした検索が可能である。

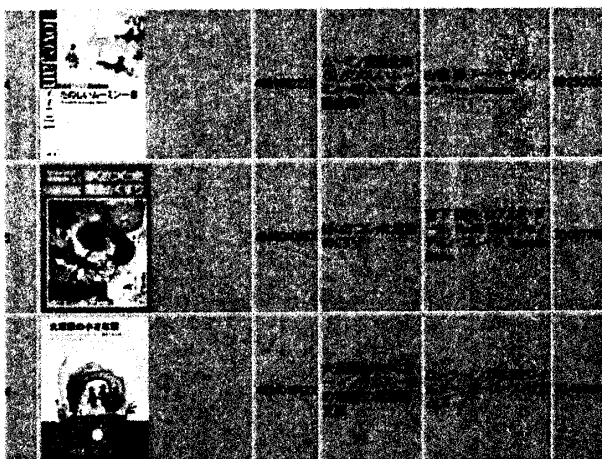
- 2) 日本の児童図書・ヤングアダルト図書に対する感性語の付与を行った先行研究⁷⁻¹⁰⁾を元にして、検索用8項目、入力用12項目の感性語項目を設定した。
- 3) whichbook.netと同様に専門家による感性項目ごとの値を入力するシステムを持つが、それ以外に初期データとして書評データを元にして各図書に対する感性語項目への値を設定している。書評データを元にした感性語項目の値は、あらかじめ人手によって529個の形容詞について12個の感性語項目との対応関係を設定しておき、それを利用した。なお、書評中に対応関係が設定された形容詞が複数出現した場合は、その感性語項目には設定された値の平均値を付与した。
- 4) whichbook.netは検索時に4つまでの感性語項目の値を入力し、その入力された感性語項目の値と図書に付与された値との相違を元にした検索を行っていたが、このシステムでは検索用の項目と入力用項目とのベクトル空間上での距離を計算して近い順に出力するという形式を持っている。

作成したシステムの検索画面と検索結果表示画面を第1図および第2図に示す。

第1図 検索画面

空しい 你: 空しい	1 2 3 4 5	思ったとおり 你: 思ったとおり
あたたかい 你: あたたかい	1 2 3 4 5	孤独な 你: 孤独な
夢の中の 你: 夢の中の	1 2 3 4 5	現実世界 你: 現実世界
驚かれる 你: 驚かれる	1 2 3 4 5	心が震える 你: 心が震える
心がじんとする 你: 心がじんとする	1 2 3 4 5	しめられた 你: しめられた
元気が出る 你: 元気が出る	1 2 3 4 5	がっかりする 你: がっかりする
楽しい 你: 楽しい	1 2 3 4 5	悲しい 你: 悲しい
寂しい 你: 寂しい	1 2 3 4 5	辛い 你: 辛い

第2図 検索結果画面



このシステムを用いた検索実験では、おおむね人間の判断と近い図書が出力され、その有効性が確認された。なかでも、専門家が人手によって感性語項目に対する値の付与を行った図書については高い評価が得られた³⁾。それに比較して、書評を元に感性語項目の値を付与した場合は、比較的评价が低かった³⁾。これは、書評を元にした場合には、入力用に設定した12個の感性語項目のうち平均2.53個の項目のみにしか値が付与されなかったため、識別力がそれほど高くなかったためであると考えられる。そこで、本研究では専門家が入力したデータを元にして、多くの感性語項目についても適切な値を自動的に付与することができる仕組みを開発した。

2. 感性語項目に対する値の付与

作成したシステムで、書評中の語を元に感性語項目を付与することが困難であるのは、書評中の語のうち、明確に感性語項目として設定するために使用できる語が529個と多く

なかったことが大きな原因であった。

そこで、本研究では図書の書評中に出現する語のうち、直接的には感性語項目を示すものではない語についても、専門家が入力した数値を元に感性語項目との対応関係を与えることとした。すなわち、専門家が感性語項目を付与した図書の書評中に出現した単語のうち、ストップワードを除く語を対象に、その語が図書の書評中に存在した場合に、各感性語項目に対してどのような値を付与することになるのか(または、その感性語項目には関係ないのか)を、専門家が付与した最適解を元に決定しようとするものである。

このような方法を考える場合、最適解が付与された図書の書評中に存在する全ての語が感性語項目と関連しているわけではなく、関係のない語に対して誤った値を設定してしまうという問題がある。すなわち、あまり重要ではない語の共出現と、より重要と思われる語の共出現が、類似度を計算する上で同じ重みを持つことになる。しかし、専門家が感性項目に値を付与した図書と新たに感性項目に値を付与しようとする図書とが、その書評中に同一の語が含まれているとした場合でも、その語の持つ意味によって、感性項目に値を付与する上での有効性には大きな違いがあると考えるのは極めて自然である。

本研究では、この問題に対して、各語に対する出現頻度および、他の感性語項目との関連度の強さに基づいて設定した基準をクリアする語のみを対象とした分析を行うこと、また、語ごとに重みを設定することで解決している。

前者は、語を元に感性語項目の値を設定する影響力が、ある語が対象とする図書の書評中に出現する総回数が多くなるほど小さくなり、また、他の感性語項目の値を決定する影響力が大きい(たとえば1から5の5件法で値を付与する場合には3からの離れた値を付与する力が大きいほど影響力が大きい)ほど小さくなると考え、閾値を設定して、この値が大きい語のみを対象としたものである。

後者は、丸山らがマイクロフィーチャーを用いた類似度の計算において提案している、用いる観点に重みをつけることで構成された空間をゆがめて、より人間の直感に適合する方法を適用するものである¹¹⁾。すなわち、そこで、最急降下法を用いて最適解と計算値の誤差の2乗が小さくなる方向を計算することで重みを変化させる手法である。最急降下法にもとづいて重みを変化させ、それによって望みの関数を求めるという手法には、ニューラルネットワークにおける誤差逆伝播法(バックプロパゲーション)がある。本研究では、この最急降下法を用いたこの手法によって、書評中に出現する各語に付与する重みを変化させることとする。本研究では、 $2n$ 個の入力と1つの出力を持つニューラルネットワークを考えることとした。

具体的には、実際に計算された感性項目に対する計算値と、専門家によって与えられた感性項目に対する値の差を誤差と考える。この誤差を、書評中の語に付与された重みの値に応じた関数(誤差関数)とみなす。各語に付与された重みの値を w_i 、誤差関数を $g(w_i)$ とすると、 w_i の最適値は誤差が最少となった時の値であると考えられる。この最適値を求めるために、書評中の語を元に感性項目の値を決定する式を、各語に付与した重みで偏微分する。この微分係数を用いて誤差関数の極小値を求めることができる。この方法を使うことにより、 w_i を順に少しずつ変化させて最小点を求める方法に比べて、より速いスピードで収束させることができる。すなわち、語に付与された重みで感性項目に対する値が変化するとき、重みで値を微分することによって、値を希望の数字にするために必要な重み

の変化量の最適値を求めることができる。

3. 付与実験と今後の課題

専門家5名が最適解を付与した図書は合計で215冊であった。215冊の図書のうち、書評が存在していた126冊を元にして、形容詞等に対する重みの付与を行った。

126冊の図書の書評中の形容詞数はのべ772語(平均6.1語)、名詞は4470語(平均35.5語)であった。これらの形容詞および名詞を用いて付与実験を行ったところ、従来は平均2.53個の感性語項目にしか値を付与することができなかったものが、平均5.35個の感性語項目に値を付与することが可能となった。その際、形容詞だけではなく一部の名詞についても感性語項目の値を決定する上で有効に用いることができた。

このことは、専門家が付与した最適解を元にして、単語の意味そのものからは感性語項目に対する値の付与に用いることが困難であった名詞や形容詞などについても感性語項目の値を決定するのに有効であることを示していると考えられる。

今後は、本手法で付与した感性語項目が専門家の判断と比較してどの程度妥当な結果であるかの検討が必要であると考えられる。また、本研究で用いた手法はあくまで距離計算の一例であり、この手法が最も適切かどうかの検討が必要となろう。

4. 注・引用文献

- 1) 桑田てるみ. 英国の読書検索サイト"whichbook.net"のわが国における意義と可能性: 学校図書館員が読書案内に利用する視点から. 2004年度三田図書館情報学会研究大会発表論文集. p. 5-8(2004)
- 2) whichbook.net. [2005. 12. 10]<<http://www.whichbook.net/index.jsp>>
- 3) 桑田てるみ;原田隆史;小林加奈.感性語を検索語とした読書資料検索システムの開発に関する基礎研究-日本語版 which book の作成-. 学校図書館学研究. Vol.8 (2006) 投稿中. また, 作成中のシステムは <http://whichbook.slis.keio.ac.jp/> にて公開している。
- 4) アマゾンジャパン. Amazon.co.jpへようこそ [2005. 12. 18]<<http://www.amazon.co.jp/>>
- 5) ヤングアダルト図書総目録刊行会. ヤングアダルト図書総目録 YA BOOKS CATALOG 2004. ヤングアダルト図書総目録刊行会, 2004, 314p.
- 6) ビーケーワン. オンライン書店bk1. [2005. 12. 18]<<http://www.bk1.co.jp>>
- 7) 三和義秀;小林久恵. 小説を対象にした感性語の分類の基礎研究: 意味的類似性を基準として. Journal of Library and Information Science. Vol. 17, p. 27-37(2003)
- 8) 原田隆史. 書評中の感性語を用いた児童書・ヤングアダルト図書の自動分類. 生涯学習時代における学校図書館パワー(渡辺先生古希記念論文集). 2005, p. 225-246.
- 9) 原田隆史. 書評中の感性キーワードを用いた小説の分類. 情報知識学会誌. Vol. 15, No. 2. p. 51-54(2005)
- 10) 小林久恵. 読書感分析に基づく情報検索インターフェースの研究: 読書感分析システムを中心に. 中央大学大学院社会情報学専攻修士論文. 2001
- 11) 丸山宏ほか. マイクロフィーチャーに基づく語問の関連度の計算とその適合化. 電子情報通信学会全国大会予稿集. 1989-05(NLC89-1) p. 1-8(1989)

新たな文脈を生成するデジタルコンテンツ構成法

○宇陀則彦, 松村 敦, 村田典子

筑波大学大学院 図書館情報メディア研究科

A Method of Reconstruction of Digital Contents

Norihiko UDA, Atsushi MATSUMURA and Noriko MURATA

This paper describes a method of reconstruction of digital contents. Recently, a lot of libraries and museums digitize documents or objects, and construct digital archival systems. However, most of them do not make an advantage of digital contents. We take context of annuals of Asian affairs to pieces, and rearrange the pieces by relating same subject between headlines and texts. As a result, we showed the method generated new contexts which the documents have originally.

1. はじめに

近年、多くの図書館、美術館、博物館が資料の電子化を進めている。しかしながら、電子化されたコンテンツは検索が可能になったという点以外、電子化したことのメリットが活かされていない。そこで本研究では、自由に組み替え・並べ替えが可能であるというデジタルコンテンツの特徴を活かし、紙媒体の資料では困難であった多様な視点からのアクセスが可能になることを示す。題材としてアジア動向年報[1]を取り上げる。

2. アジア動向年報の構造

アジア動向年報の作成元であるアジア経済研究所は、アジア地域の経済、政治、社会の諸問題に関する基礎的・総合的研究とその情報提供を行う機関である。アジア動向年報はアジア各国の政治・経済・社会動向の概観と分析、統計データ、地理情報などに加え、現地新聞の記事を基に作成された重要日誌で構成されている。1冊の中では、まず『総論』・『主要経済指標』・『主要トピックス』・『各国・地域の動向』という大見出しに文書が分けられている。重要日誌は日付と共にその日に起こった出来事が1文程度で記述されている。

重要日誌

1月1日、『人民日報』元旦社説が世界の多極化と経済のグローバル化を強調
1月12日、昨年度中米貿易総額は549億4000万ドルで前年比12%増。中国の対米貿易黒字は210億2000万ドル
1月13日、1998年の中国の小売物価が前年より2.6%低下
1月14日、中国のインターネット利用者が150万人を突破

図1：重要日誌（1999年度版中国の重要日誌から抜粋）

一方、本文では、ポイントとなる出来事についての詳細、見解等が記述されている。本文中で記述されている出来事の多くは重要日誌にも記述されており、内容が重複する重要日誌と本文の一部では、ほぼ同様の語句が使用されている（図2）。このことから、重要日誌に記述されている事項・事件の中でも、特に重要な事項・事件が本文中で詳細に述べられていると考えてよい。

重要日誌：1月1日、『人民日報』元旦社説が世界の多極化と経済のグローバル化を強調。

本文：1月1日 付けの中国共産党機関紙『人民日報』は社説で世界の多極化と経済のグローバル化を強調した。当時の中央指導部の関心は、前年末までに明らかになった経済の減速傾向に歯止めをかけることにあった。なかでも中国は外資の動向には神経質にならざるを得なかった。

図2：重要日誌と本文に共通して出現する語句の例

本研究はこれらのことを踏まえ、アジア動向年報の文書構造を解体し、内容に沿って並べ替えることにより、コンテンツの再構成を行った。重要日誌と本文に共通する語句を介して両方の記述を関連付けることにより、重要日誌と本文を内容ごとに整理する。さらに、整理された重要日誌と本文を時系列に並べ替えることによって、新しい文脈を生成する。

3. アジア動向年報の再構成手法

3.1 解体

① 本文の解体

内容ごとのまとまりは段落単位で記述されている。そこで、アジア動向年報の本文を段落の単位に解体し、内容ごとに分類した。そのうえで、のちほど関連付けの処理をしやすいように図3のようなXML形式に手作業で変換した。

```
<DOC NAME="china1999about1">
<TITLE>概況 1</TITLE>
<TEXT>
江沢民国家主席の言葉を借りれば、「1999年は実に非凡な年」であった。中華人民共和国は10月1日建国50周年を祝った。記念式典は「3代の指導集団の下での新中国の社会主義革命・建設事業の輝かしい成果」を誇示し、天安門事件以来10年に及ぶ江沢民政権の安定を内外にアピールするものであった。11月には中国はアメリカとWTOへの早期加盟について合意し、12月にはマカオの返還を達成した。
</TEXT>
</DOC>
```

図3：段落の単位へ分解

分解された段落は、<TEXT>タグで囲まれている部分である。<DOC>タグの属性NAMEは、各段落を識別するための名前を指定し、<TITLE>タグで囲まれている部分では各段落のタイトルを指定した。タイトルはその段落の属する節の見出しに段落番号をつけたものである。

② 重要日誌の解体

重要日誌は日毎に起こった出来事について1行に1文程度で記述されている。内容ごとに分けるために、重要日誌は行の単位(1日分)に解体する。重要日誌のデータはテキスト形式で、1日分が1行で記述されている。よって、ファイルから1行ずつ読み込んで、行の単位で抽出した。

3.2 関連付け

重要日誌はその日に起こった重要項目・事件について濃縮された内容となっていることから、重要日誌に出現する名詞句は、どれも内容を表す上で重要な語句である可能性が高い。一方、本文は主題構造を持つと仮定し、文中の名詞句は主題語とそれ以外のものに分けた。各文において主題語はその文のテーマであり、内容を表す上で重要な語である。よって、重要日誌に含まれる名詞句と段落の各文に含まれる主題語との共通性を見出すことで重要日誌の各文と本文中の段落を関連付ける。関連付けのフローを図4に示す。

① 重要日誌の抽出

関連付けの起点となる重要日誌の各行から日誌ナンバー、国名、年月日を取り除き、後に抽出する名詞句の識別に使用するため、日誌ナンバーを記録する。

② 関連付ける段落の絞込み

①で抽出された各重要日誌に対して関連付ける候補の段落を絞り込む。重要日誌を質問文として検索を行い、関連度の高い段落をまとめて同一ファイルに保存する。

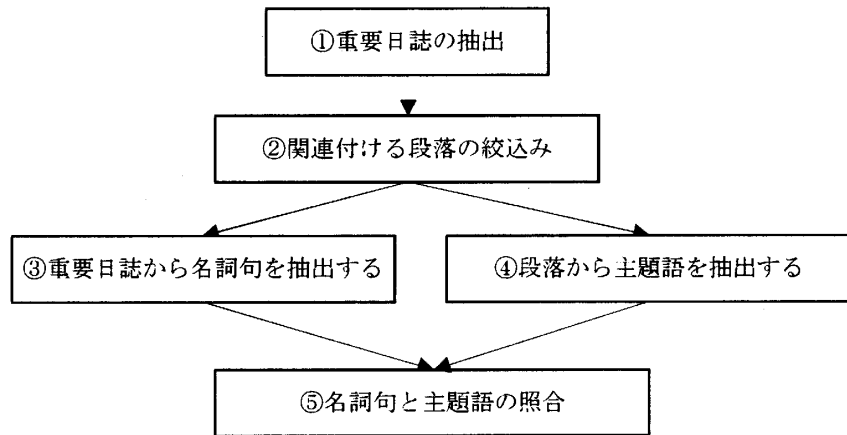


図 4：関連付けのプロセス

③ 名詞句の抽出

重要日誌に含まれる名詞句と段落の各文に含まれる主題語の照合を行い、一致すれば関連付けを行う。具体的には、形態素解析システム茶釜を使用して名詞句を抽出した。抽出された名詞句は、どの重要日誌から抽出されたものが明確にするため、抽出元である重要日誌と同じ日誌ナンバーを付けてファイルに保存した。

④ 主題語の抽出

②の処理で検索結果のスコアが最も高かった段落は、該当する重要日誌との関連付けが成立する可能性が最も高い段落であると考えられる。そのため、まずスコアが最も高い段落を文の単位に解体し、ここから「は」と「が」格を主題語として抽出した。さらに、英字のみからなる語も主題語として抽出した。

⑤ 名詞句と主題語の照合

②で行った絞込みによって、ある程度内容が関連した段落が関連付けの対象候補となっている。よって、重要日誌から抽出された名詞句と、段落から抽出された主題語を順に照らし合わせ、一つでも一致すれば関連付けを行う。

3.3 テーマ抽出

関連付けの処理で抽出された名詞句と主題語は、それぞれ重要日誌と段落の内容を表す語句である。これに対しテーマとは、利用者がアジア動向年報にアクセスするための視点である。したがって、アクセスするうえで重要かつ一般的によく知られた語句を抽出する必要がある。そこで、本文と重要日誌両方に出現する名詞句をテーマとして抽出した。

① 語句の抽出

テーマとして抽出する語句は、そのテーマに沿って再構成された文書の見出しとなる。まず、茶釜を使用して、各重要日誌と本文を形態素解析し、形態素の単位で抽出する。ただし、括弧で囲まれた語句・文については有名な事項・事件の名称、新聞・雑誌名、記事などが入っているため、括弧の中身だけを抽出する。

② 語句の絞込み

①で抽出された語句を見ると、固有名詞、英字のみからなる語、括弧で括られた語句・文が、比較的知名度が高いことが分かった。そこで、茶釜の解析結果に含まれる品詞情報を利用して、①で抽出された語句の中からさらに絞り込んで固有名詞、英字のみからなる語、括弧で括られた語句・文を抽出した。

③ 照合と抽出

テーマとして抽出されるのは、重要日誌と本文両方に存在する語句である。よって、重要日誌から抽出された語句のリストと本文から抽出した語句のリストを照合して、両方のリストに存在する語句のみを抽出する。

3.4 並べ替え

抽出されたテーマに沿って重要日誌と段落を並べ替え、新たな文脈を生成する。各重要日誌には日付順にナンバーが付いているという特性を活かし、集められた文と段落をナンバー順（日付順）に並べ替えた。また、複合語をテーマに指定したい閲覧者のために、複数のテーマを組み合わせた並べ替えも試みた。ただし、今回は複数のテーマが出現する重要日誌が少なかったため、選択されたテーマのいずれかが出現する重要日誌とその重要日誌に関連付けられた段落を集めて並べ替えるにとどめた。

4. アジア動向年報デジタルライブラリシステム

本研究はアジア動向年報 1999 年度版中国を対象に、オリジナル文書と再構成された文書の両方でアクセスできるプロトタイプシステムを構築した。画面左に国名、画面下部に年を表示することで、オリジナル文書と同じアクセスを行いたい利用者に配慮した。一方、別の観点からアクセスしたい利用者のために、画面上部に主要トピックスへのメニューを配置した。

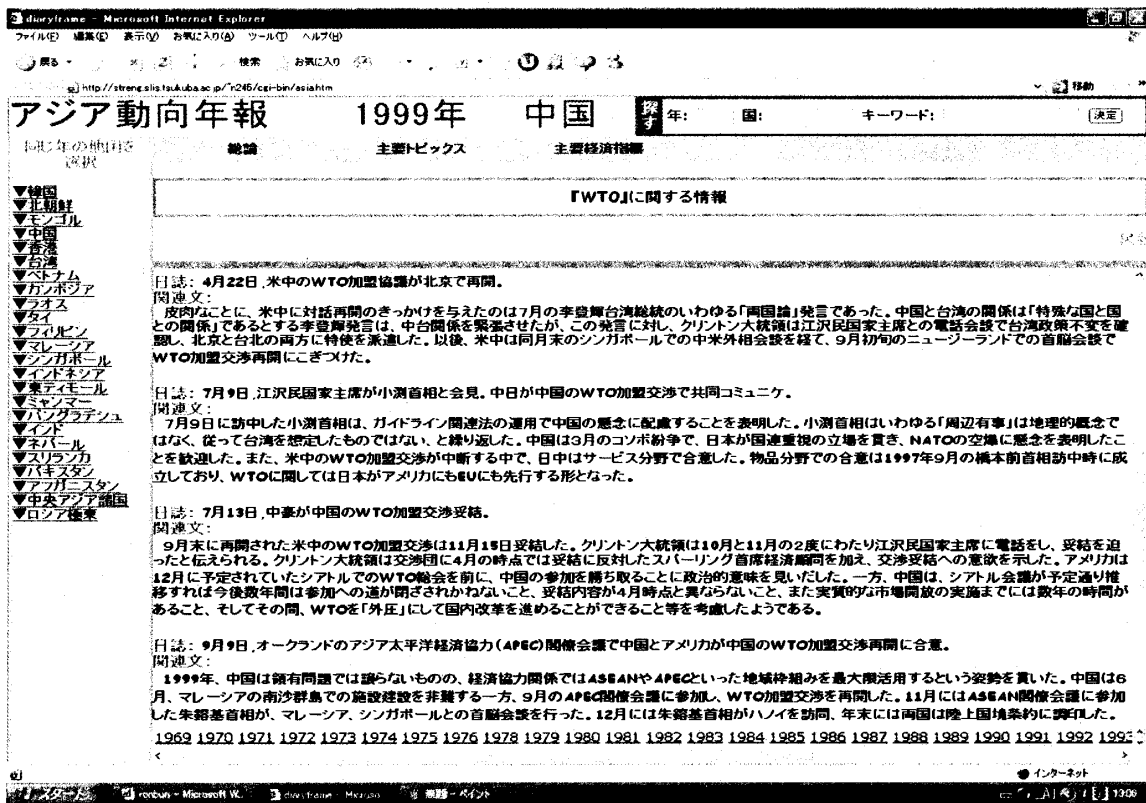


図 5：再構成された文書

アジア動向年報から抽出されたテーマから一つを選択すると、そのテーマに沿って重要日誌と段落がまとめられ、時系列に並べ替えられて、再構成された文書として表示される（図 5）。表 1 に新しく追加されたテーマ一覧を示す。

表 1：新しく追加されたテーマ一覧

種類	テーマ
人名	江沢民, 鎔基, エリツイン, 小平, クリントン, 石原, シュレーダー, 金永南, 毛沢東, 小淵
国・地域名	マカオ, 台湾, 福建, 朝鮮民主主義人民共和国, 北朝鮮, 江蘇, マケドニア, 北京, 湖北, ロシア, ユーゴ, アメリカ, カナダ, アジア, 東京, 欧州, 四川, 南海, ドイツ, 香港, 上海, 日本, コソボ, マレーシア, 南沙, パプアニューギニア, 中日, シンガポール, キルギス, カザフスタン, タジキスタン, 東ティモール, 南京, ニューージーランド, ワシントン, 天津, 中東, インドネシア, ハノイ, シアトル, チェチェン
新聞・雑誌名	人民日報, 解放軍報
組織名	WTO, ASEAN, NATO, 下院, APEC, 法輪功
その他	行長, GDP, 革命烈士

図 6 は既存の文書と再構成された文書を比較したものである。既存の文書では、WTO と NATO に関する重要日誌や段落が、4つの節「国内政治」「経済」「対外関係」「重要日誌」に分散している。そのため、WTO についてまとまった情報を得るためには、複数の節を移動しなければならない。一方、再構成された文書では、既存の文書においてばらばらになっていた WTO と NATO に関する内容の重要日誌と段落が、それぞれまとめられている。再構成された文書を利用することによって、複数の見出しを移動しなくても、特定の事項・事件についてまとまった情報を得ることが可能になった。

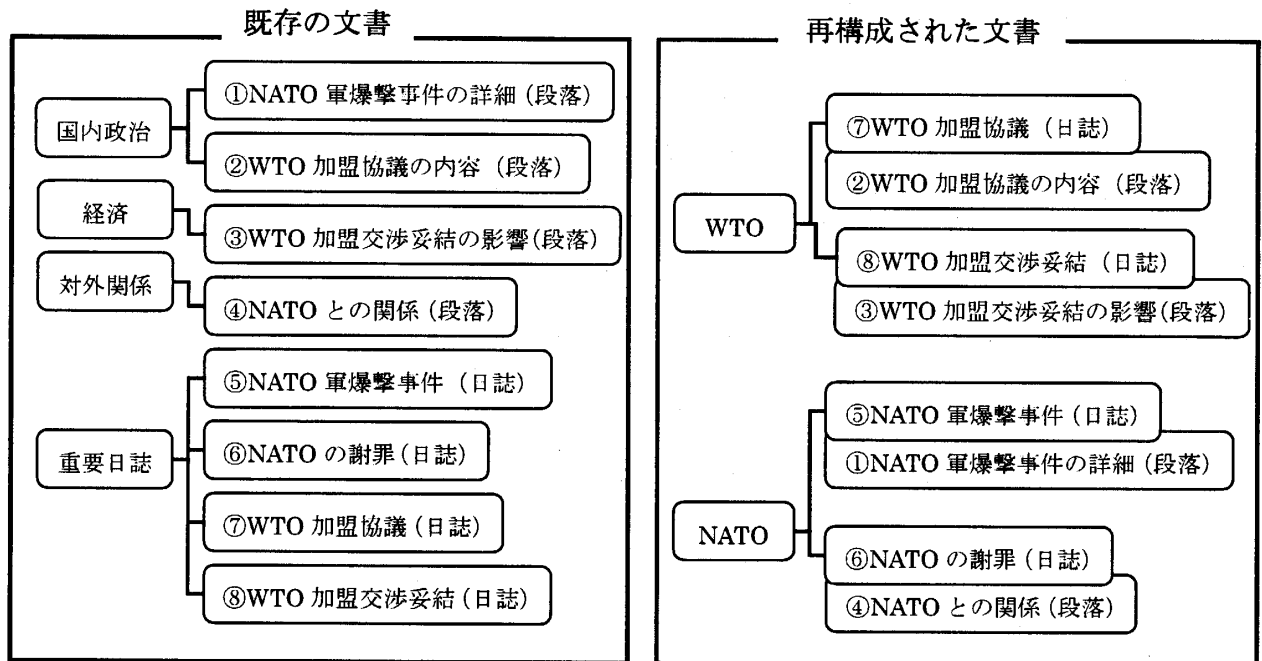


図 6：既存の文書と再構成された文書の比較

5. 考察

- 解体について

解体の処理では、本文を段落の単位、重要日誌を行の単位（一日分ごと）に解体した。

解体された段落を見ると、そのほとんどが内容ごとに分けられているため、この手法は有効であったと考えられる。さらに、一般的に段落は一つの内容を持ってまとまった部分のことを指す。このことから、アジア動向年報以外の一般的な資料に対してもこの手法を用いて解体を行うことが有効であると考えられる。しかしながら、段落がどの程度詳細な内容に分けられているのかは、個々の資料に依存する。よって、各段落の内容の詳細さを把握した上で、何の単位に解体するのが妥当かを検討する必要がある。

- 関連付けについて

関連付けの処理では、解体された重要日誌と段落の関連性を、重要日誌に含まれるすべての名詞句と、段落に含まれる主題語の共通性によって定めた。この基準は、重要日誌に出現する語句のほとんどが重要である可能性が高く、本文では重要な語句が占める割合は低いという性質によるものである。

しかし、本手法を用いた結果、関連のある段落が関連付けの対象候補になっているにも拘らず、名詞句と主題語が一致しないために関連付けが成立しないまたは別の段落と関連付けが成立してしまうことが多くある。今後は、本手法による関連付けの再現率について評価し、より高い精度で関連付けを行うための手法について検討する必要がある[2]。

- テーマ抽出について

テーマ抽出では、人物名、国・地域名、新聞・雑誌名、記事、組織名等が抽出され、比較的知名度が高い語句が抽出された。また、抽出された64のテーマ中、理解できるテーマは58あり、適合率は高いといえる。

しかし本手法には、重要日誌にのみ出現する語句、本文にのみ出現する語句がテーマとしてふさわしくないのかということについて検討が不十分である。また、本研究では固有名詞など知名度の高い語句を中心にテーマを抽出したが、調べ物をする利用者にとっては知名度の低い語句もテーマとして提示されている方が便利であると考えられる。今後は、本手法によってどの程度網羅的にテーマが抽出できたかを明らかにするために評価実験を行い、適当なテーマ抽出方法について検討する必要がある。

- 並べ替えについて

本研究では、重要日誌の各文に含まれている日付情報と日誌ナンバーに着目し、重要日誌とその重要日誌に関連付けられた段落を時系列に並べ替えた。アジア動向年報でなくても、重要日誌のように、時系列に整理されている見出しや段落などのデータが含まれている資料であれば、この手法を適用し、推移・経過の文脈を作り出すことが可能である。

6. まとめ

本研究では、アジア動向年報を対象としてデジタルコンテンツの再構成を行い、プロトタイプシステムを構築した。本システムによって閲覧者は多様な視点からアジア動向年報のコンテンツへアクセスすることが可能となった。今後はより広い範囲のアジア動向年報や他の電子化された資料に対して本手法を適用し、その有効性を検証する予定である。また、他のコンテンツの再構成に関する研究[3]との比較も行いたい。

謝辞

本研究はアジア経済研究所図書館の協力の下に行われました。ここに記して感謝の意を表します。

参考文献

- [1] 鳥谷尾克男ほか編. アジア動向年報. 千葉, アジア経済研究所, 2001. 636p.(ISBN4-258-01001-4)
- [2] 乾孝司, 奥村学. 因果関係タグつきコーパスの構築と分析. 言語処理学会第11回大会発表論文集, 2005, p.486-489.
- [3] 北憲一郎, 吉場慎二, 富樫雅文, 大岩元. TransPublishingによる高校情報教科書の再構成. 情報処理学会研究報告 コンピュータと教育. 55(5), 2000, p.33-40.

問題解決を目的としたWWW情報探索行動の特徴

Features of information seeking behaviors on the WWW for problem solving

相良佳弘(聖徳大学)

SAGARA, Yoshihiro(Seitoku University)

Information seeking behavior on the WWW attract public attention as the WWW contents came to be used widely as means for problem solving. The object of this study was to investigate features of information seeking behaviors on the WWW for problem solving. Ten subjects were college woman, and they had the problem of job hunting. Findings indicated that the WWW users had a tendency to pick up relevant pages from plenty of search results which were retrieved by simple query, and to value the pages from the search engines.

1.はじめに

今日、インターネットは日常生活において幅広く用いられるようになってきている。特にティーンエイジャーから 30 代における利用率は高く、この世代の情報利用者にとって主要な情報源の一つといえる。それに伴い、WWW を対象とする情報探索行動に関心が高まり、これまでも様々な研究が行われてきた。本研究では、利用者自身が抱える問題状況を解決するための手段として WWW での情報探索を捉えた場合、どのような特徴があるのかを検討する。

2.研究の背景

2.1 WWW における情報探索

WWW は非常に身近であると同時に、膨大な情報を蓄積しているメディアである。これらを効率的に利用できるよう、検索システムのユーザビリティの向上や検索手法の改良とともに、利用者の検索行動、探索行動に関する研究も行われてきた。

Jansen 等⁽¹⁾ は、WWW 情報探索を対象とした先行研究をレビューし、WWW の情報探索には、1) 1つの検索式あたりキーワードを 2 語程度しか用いない、2) 1回の探索あたり用いる検索式の数が少ない、3) 検索結果を最初の 10 件程度しか閲覧しないといった特徴があると述べている。従来の書誌データベースの検索に比べると用いられるキーワードの数が少なく、検索式の数も少ない。また OPAC を用いた情報探索では、しばしば多くの件数を出力する事例が見られるが、WWW 情報の探索では少ない件数しか出力されないという特徴がある。このように WWW 情報探索では、従来の書誌データベースや OPAC とは異なった特徴がみられ、その背景には WWW 情報の特性があると考えられている。

齊藤等⁽²⁾ は、WWW 情報探索を問題解決活動としてとらえ、科学的発見のモデルからその認知プロセスや利用者の知識、経験がおよぼす影響について明らかにすることを試みた。その結果、知識や経験によって検索式やキーワードの種類数に差異は見られなかった

ものの、経験豊富な利用者は少ない検索結果の閲覧から適切な情報を得ていることを指摘した。また、内容の質や信頼性が必ずしも保証されないという WWW 情報の特性から、種市等⁽³⁾は情報探索の中で「情報を評価する」点に着目しその行動を分析した。さらに Tombros 等⁽⁴⁾は、利用者が Web ページを評価するにあたってどのような基準を用いたかについて実験を行った。これらの研究に共通している点として、被験者に何らかの課題を与え、その課題を解決するためのプロセスを調査対象としていることがあげられる。しかしながら、これらの課題のほとんどは調査対象者自身が抱える問題状況から由来する課題、あるいは情報要求とはいえない。

2.2 問題解決のための情報探索

日常生活の中で、人間が何らかの問題状況に直面し、その解決のために情報が必要となることがあり、その結果として情報探索が行われると考えられる。下村等⁽⁵⁾は、大学生が就職活動をするにあたりどのような情報源を用いているのかについて調査を行っている。大学生にとって就職活動は、自ら解決する必要のある問題状況であり、その中で企業の説明会や友人、OB・OG、WWW 上の各種情報などの情報源を利用する。この調査の結果、企業のホームページや就職活動に関するホームページ⁽⁶⁾は誰でも利用可能であり、就職活動の初期段階において重視される情報源であることが示された。

問題状況を抱えた利用者は Taylor⁽⁷⁾の情報要求形成の四段階に示されるように、明確にはない直感的要求を徐々に意識するようになり、それを他者に伝えられるよう表現し、さらに検索システム等のルールにあわせた形式で表していくと考えられる。この時、問題状況を解決する情報が事実や既知事項の確認であるならば、サーチエンジンを利用するにあたってその既知事項を含むデータ検索を行えば十分であり、客観的な正しさの確認も容易である。しかしながら利用者が抱える問題状況には、前述の就職活動のように未知の事項を探索する事例もある。このような未知事項の探索においては、何について検索するのか、また検索に用いるキーワードや検索式、そして得られた情報の妥当性を判定することが困難である。こういった状況下で、どのような情報探索が行われているかについて明らかにすることは大きな意義があると思われる。相良⁽⁸⁾は、学部3年生を対象に「就職活動」に関する情報を探索させる実験をし、同じ問題状況を抱える10名の被験者が、どのような主題を検索し、どのような情報を適合と判定するかについて分析を行った。その結果、同じ問題状況を抱える被験者間で異なる主題を求めていた事例がみられ、いわゆるデータ検索的な検索式が用いられていたとしても、その背景にある認知的要求に基づいて適合性が判定されることが示された。本研究ではこの「就職活動」に関する情報探索にどのような特徴があったのか、さらなる検討を行う。

3.被験者自身が抱える問題解決のために行われる情報探索の特徴

3.1.対象

本研究では、相良⁽⁸⁾の実験で得られたデータを検討の対象とする。実験は、学部3年生が就職活動の準備を開始する時期にあわせ、2005年6月末から7月半ばに行われた。10人の被験者はいずれも聖徳大学人文学部に所属する女子大生であり、就職活動のために何かしなければならぬという社会的な状況におかれている。こうした状況下の被験者に対

して、「就職活動をするにあたって、どのようなことに気をつけなければいけないのか、注意しなければいけないのか知る」必要があるという問題状況を提示し、その問題状況に対して WWW 上の情報を探索するよう指示を与えた。被験者は自由に WWW コンテンツを探索し情報収集を行い、そのプロセスをデジタルビデオに収録した。その際、被験者自身が役に立つと感じたページの URL をブラウザの「お気に入り」機能を用いて保存するように求めた。さらに探索終了後、情報収集の結果に対する満足度や、普段と同じように検索できたかなどをアンケートおよびインタビューによって調査した。

3.2.結果

今回の実験では、サーチエンジンを用いることを指定しなかったが、全ての被験者が利用していた。被験者ごとの検索回数、平均キーワード数等を表1に示す。ほぼ全ての被験者が2語前後のキーワードからなる検索式を用いており、利用者自身の問題状況に基づく探索においても従来の研究と同様に、比較的単純な検索式が用いられていることが明らかになった。しかしながら4名の被験者が、4回以下の検索回数で探索を終了していた一方で、10回以上の検索を行った被験者も見られた。表示ページ数は、サーチエンジンを利用した際に検索結果を何ページ目まで表示したかである。平均値をとった場合、全ての被験者が1から2ページ前後の検索結果を閲覧していることになるが、個別にみた場合、被験者によっては7ページ目140件（被験者G）、あるいは6ページ目60件（被験者C）まで閲覧した検索事例もあった。表示件数は、検索結果のリストを何件分閲覧したかを示す値である。探索を終了するまでに行われた検索回数が多いことから、100件以上出力した被験者が多かった。実際には検索結果以外にもスポンサー、広告等のリンクが表示され、サーチエンジンによってはカテゴリ等へのリンクも示されるため、これ以上のリンクを被験者は閲覧したと考えられる。

表1 被験者ごとに見る検索回数、キーワード数

	A	B	C	D	E	F	G	H	I	J
検索回数	11	13	3	2	4	37	14	1	14	7
平均キーワード数	2.0	1.9	2.0	2.0	1.5	2.2	1.6	2.0	1.6	2.1
表示ページ数	1.3	1.1	2.7	2.0	2.0	1.3	2.0	1.0	1.6	1.0
表示件数	200	130	90	50	100	476	350	10	230	110
クリック数	39	22	25	13	29	60	40	5	24	20
有用ページの得られた結果	13	5	7	5	5	5	6	1	12	2
有用ページ数	14	11	14	5	7	6	6	1	14	2
有用サイト数	13	5	9	5	5	5	6	1	13	2

これら表示された結果のうち、スポンサーリンクやカテゴリ等へのリンクも含め、被験者が選択した数がクリック数である。また、このうち1件以上の有用ページにつながった数を有用ページの得られた結果に示す。全体で64件のサイトが被験者によって有用と判定され、その中でサーチエンジンの検索結果に含まれていなかったサイトは4件だけであった。このうち3件は、検索結果に含まれていたサイトのリンク集から発見された。ページ単位では必ずしもサーチエンジンが提示したページが有用と判定されたわけではないが、サイト単位でみた場合、ほとんどの有用サイトが検索結果に含まれていた。一方、閲覧したサイトのリンク集や外部リンクから有用サイトが得られた事例は少なかった。

なお、情報収集の結果に対する被験者の満足度は概ね高く、不満を示した被験者はいな

かった。また普段と同様の情報収集が行えたかを 5 段階で尋ねた質問に対して、4.7 という高い値が得られ、被験者自身が普段行っているのとほぼ同様な探索が行われたことが示された。

4. 考察

有用と判定されたサイトの多くがサーチエンジンから得られていたことは、サーチエンジンが WWW 上の情報探索に果たす役割が大きいことを示すものと考えられる。一方で、リンク集や他サイトへのリンクはあまり活用されていない傾向が見られた。検索式に用いられるキーワード数は先行研究と同様に少なく、単純な検索式が用いられていることが確認できた。また、検索回数や表示件数は被験者によりばらつきが見られた。検索回数の多い被験者は、得られた検索結果や Web ページからキーワードを拾い出しており、暫定的な検索式での検索を繰り返したと考えられる。書誌データベースの検索のように複雑な検索式を用いて検索結果を絞り込むのではなく、多少のノイズを想定した上で単純な検索を行い、多くの出力件数の中から選択していると思われる。このことは、表示された件数のうちクリックされた件数が 2 割に達していないことからもいえよう。

本研究では、被験者自身の問題状況に基づく情報探索にどのような特徴があるのかについて検討した。利用者自身にとって未知の状況に対する情報探索の場合、どの情報が正しいのか、あるいはどの情報が適切なのかを判断することが困難である。しかしながら被験者は情報を探索し、その結果に満足したと答えている。どのような情報が得られた場合に利用者は満足するのか、あるいは意識されない要求からどのような表現された要求が生成されるのかについて今後、さらなる検討を加えていく予定である。

(1) Jansen,B.J. ; Pooch,U.: "A review of web searching studies and a framework for future research", *Journal of the American Society for Information Science and Technology*. Vol.52, No.3, pp235-246, 2001

(2) 齋藤ひとみ ; 三輪和久 : 「問題解決行動としての WWW 情報探索 : 科学的発見の枠組みに基づく検討」, *Cognitive studies*, Vol.10, No.2, pp258-275, 2003.

(3) 種市淳子 ; 逸村裕 : 「サーチエンジンにおける探索結果を評価する行動の分析」, 日本図書館情報学会, 三田図書館・情報学会合同研究大会発表要綱 2005, 東京, 2005-10, pp173-176.

(4) Tombros,A. ; Ruthven,I. ; Jose,J.M.: "How users assess webpages for information seeking", *Journal of the American Society for Information Science and Technology*., Vol.56, No.4, pp327-344, 2005.

(5) 下村英雄 ; 堀洋元 : 「大学生の就職活動における情報探索行動 : 情報源の影響に関する検討」, *社会心理学研究*, Vol20, No.2, pp93-105, 2004.

(6) リクルート社の提供する「リクナビ」(<http://www.rikunabi2007.com/>) など

(7) Taylor, R. S.: "Question negotiation and information seeking in libraries", *College and Research Library*, Vol.29, pp178-194, 1968.

(8) 相良佳弘 : 「WWW 情報探索過程にみるエンドユーザの適合性判定 : システム評価の観点から」 *聖徳大学研究紀要人文学部*, No.16, pp.1-8, 2005.

ユーザの検索要求に基づいた興味関心の定量的評価

Quantified Estimation Method of User's Profile
based on Search Query堀 幸雄[†]

Yukio HORI

中山 堯[‡]

Takashi NAKAYAMA

Users' personal information interests is an important for developing more personally adaptive system to search and to collect information increased. Recently, user profiling system has been used in several research system. However, evaluation approach of these systems was lack of empirical investigation using users' explicit desire. In this paper, we show the evaluation method for user profile using users' search queries as explicit desire. After the experiment, we could confirm the effectiveness of the method.

1 はじめに

これまで情報システムはどのようなユーザが利用しても同一のサービスを提供するものであった。近年、情報の量は膨大となり、ユーザの求めるサービス、利用方法は多様化している。このような背景のなか、ユーザに適した形でシステムを利用するパーソナライゼーションが注目されつつある。ユーザの情報に対する興味や要求をモデル化し、それに基づいて大量の情報の中からユーザに適した形で情報を提供するという情報フィルタリング技術などがこれにあたる。

パーソナライゼーションには、ユーザモデル、扱う情報、提示手法の3つの自由度があり、とくにユーザの情報に対する興味を的確に表現するユーザモデルが重要となる。例えばユーザの持つ興味関心を分析する研究として橘高らの研究^[5]がある。これはユーザが閲覧したページをユーザの興味のあるページと仮定し、閲覧したページをもとにユーザの興味をモデル化している。しかしこの研究ではモデル化されたユーザの興味関心の評価に、モデル化したユーザの興味関心に合うニュース文書を提示し、提示した情報に対する適合率という指標が用いられる。

ユーザの興味関心とは情報に対する要求であり、データベースシステムに与える query のような明示的な要求と提示されるまでユーザ自身も気付いていない暗黙的な情報に対する要求の2つの側面で考えることができる。先行研究における情報の提示によるユーザモデルの評価は後者の暗黙的なユーザの興味関心の評価手法である。これではユーザモデルを漠然とした情報に対する興味関心の評価したにすぎない。

本研究ではユーザの情報に対する明示的な要求を用いてユーザの興味関心を同定する方法を提案する。ユーザが検索エンジンに投入する検索キーワードはユーザの知りたい興味関心そのものを表していると仮定し、これを用いてプロファイルを評価する方法を提案する。またユーザの興味関心のモデル化には、自己組織化マップを用いて閲覧したページの特徴をマップに投影したものをを用いた。

以下、2節で先行研究におけるユーザの興味関心のモデル化手法について述べる。そして3節で本研究で提案する興味関心パターンマップの抽出手法について述べる。4節で得られた興味関心をユーザが検索エンジンに投入したキーワードで評価する方法を述べ、5節で検討すべき課題を述べ、まとめとする。

2 関連研究

ユーザの Web 閲覧時の挙動などからユーザの興味関心を抽出する手法は多数提案されている^[6]。KANSHIN^[1] や blogWatcher^[2] などは blog サイトから RSS¹ を定期的に収集し、記事の解析を行い、blog サイト(ユーザ)の興味関心パターンの分析を行なっている。しかしながらこれら研究では blog や BBS などある特定のサービスを用いたものであり、これでは該当サービス上での興味関心は把握できるものの、サービス外を含めユーザの大域的な行動を分析していないため、ユーザの興味関心を正しく把握できていないのが現状である。

またユーザの Web 閲覧時の情報行動からユーザの興味関心に関する情報を取得する研究もある。橘高らの研究^[5]ではユーザの閲覧したページ全てを興味のあるページと仮定し、閲覧したページをもとにユーザ

[†]香川大学 情報基盤センター Information Technology Center, Kagawa University horiyuki@itc.kagawa-u.ac.jp

[‡]神奈川大学理学部 Faculty of Science, Kanagawa University, nakayama@info.kanagawa-u.ac.jp

¹Rich Site Summary: Web サイトの見出しや要約などを記述した XML データ

の興味を単語ベクトルでモデル化している。この場合 Web サーバまたはプロキシサーバのログからユーザの Web 閲覧行動の分析を行なう。また閲覧したページ全てに興味があると仮定せずに、明示的な操作 (マウス操作^[10], 閲覧中の視線^[11] など) からユーザの興味関心をモデル化する方法もある。

しかしこれら研究ではモデル化されたユーザの興味関心の評価に、モデルを元に推薦した情報がユーザに適応するか否かという観点で適合率の指標が用いられる。この評価方法はユーザの情報に対する暗黙的な興味関心を含んだ形で評価される。

ユーザモデリングというように人間を系に含む技術では観測されるデータには多様の認知的要因が関係し、不規則なノイズが含まれると考えられる。ユーザプロファイルのさらなる精度向上のためにはユーザの情報に対する明示的な要求を含んで評価することが必要である。そこで本研究では従来のユーザの Web 閲覧行動などから生成されたユーザモデルをユーザ自身の明示的な情報に対する要求を用いて評価する方法を提案する。

本研究ではユーザの閲覧行動パターンが網羅的に記録された Web アクセスログからユーザの興味関心パターンを抽出し、自己組織化マップによりマップ上にモデル化する。ユーザの興味関心を 2 次元平面上に射影することで見た目にもわかりやすく抽出することが可能となる。そして得られた興味パターンマップがどの程度ユーザの興味関心を現しているのかを、ユーザが検索エンジンに投入したキーワードを用いて評価する方法を提案し、キャンパス内ネットワークでの評価実験結果について報告する。

3 興味関心マップの抽出

ここではユーザから興味関心マップを抽出する手法を述べる。各ユーザの興味関心マップを抽出するまでの全体の流れを図 1 に示す。

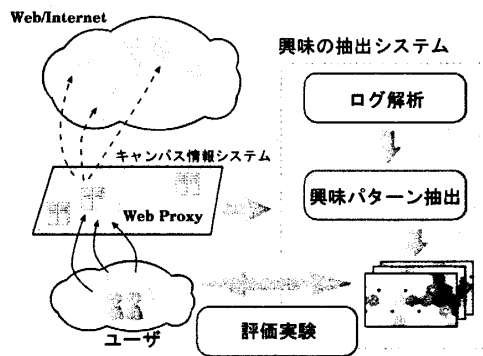


図 1: 興味関心マップ抽出システムの概要

3.1 ログ解析

まずユーザの Web 閲覧履歴を HTTP プロキシサーバ²のログから取得する。プロキシサーバはクライアントから要求のあった URL をもとに Web コンテンツを取得し、クライアントへ転送するものである。プロキシサーバのログフォーマットを下記に示す。

```
1050591018.410 364 133.92.xx.yy TCP_MISS/302 500 GET www.kagawa-u.ac.jp/ - DIRECT/ 133.92.11.23 text/html
1050591018.976 270 133.92.xx.yy TCP_MISS/200 34996 GET http://www.msn.co.jp/home.armx - DIRECT/207.46.68.11 text/html
```

上記のログから閲覧日時, Web ページ (URL), クライアント IP アドレスを取得する。クライアント IP アドレスから同時刻にそのマシンを利用したユーザを特定する。ただし個人情報保護の観点から、本研究では実際のユーザ ID をランダムな ID へと変換している。また本システムで対象とする Web コンテンツはコンテンツタイプが text/html といった単語抽出可能なもののみとした。ログ解析によりユーザ u_i ごとに閲覧した Web ページ p_i の集合 $P_i = \{p_1, p_2, \dots, p_n\}$ を作成する。

次にユーザ u_i ごとにページコンテンツの特徴を解析する。まず HTML からタグを取り除き、テキスト

²<http://http://www.squid-cache.org/>

に変換処理を行なう。そしてテキスト部分に対し、mecab³を用いて形態素解析を行ない単語に分割する。本システムでは、一般名詞、及び固有名詞に属する語を対象とする。またストップワード処理として下記のもの除外した。

1. 名詞の数、接続、記号。
2. ひらがな、カタカナ 1 文字の単語。これらは形態素解析に失敗している可能性が高いために除外した。
3. 全体で出現頻度に閾値 τ を設定する。 $\tau < 10\%$ の低頻度語を除外した。

こうして各ページの特徴ベクトル $p_i = \{w_1, w_2, \dots, w_n\}$ を得る。各単語 w_i には tf-idf により重み付けを行なった。

$$w_i(t, p) = \text{tf}(t, p) \cdot \text{idf}(t)$$

$$\text{idf}(t) = \log(N/\text{df}(t))$$

ただし、 $\text{tf}(t, p)$ は Web ページ p における単語 t の出現頻度であり、 N はユーザの閲覧した総 Web サイト数、 $\text{df}(t)$ は単語 t が 1 回以上出現する Web サイト数である。ここで Web サイトとは同一ドメインのことを意味する。

3.2 興味関心マップの作成

まず各単語 w_i を Web サイトごとの tf-idf で重み付けされた値 $v_1 \sim v_n$ とし、次式のベクトルで表す。

$$w_i = \{v_1, v_2, \dots, v_n\}$$

w_i は多次元ベクトルで構成されるためこのままではユーザの興味関心の直接的把握が困難なため、適切なクラスタリングを行ない、次元圧縮を行なう必要がある。ここで自己組織化マップ^[4]を用いてユーザの閲覧した Web ページの特徴相互の距離関係を可能な限り保持した状態でこの特徴ベクトルを 2 次元平面上に写像する。

自己組織化マップは 2 層のニューラルネットワークで構成され、入力層に入力される特徴ベクトルを w_i 、出力層のノードに連結される参照ベクトルを、

$$m_i = \{\mu_{i1}, \mu_{i2}, \dots, \mu_{in}\}$$

とする。自己組織化マップ生成の手順は次の通りである。

- 1) 参照ベクトル m_i を初期化する。
- 2) 次式により入力 w_i に対する最近傍ノード c を決定する。

$$c = \min \|m_i - w_i\|$$

- 3) 最近傍ノード c とその近傍ノードを更新する。

$$m_i(t+1) = h_{ci}(m_i(t) - w_i(t))$$

$$h_{ci} = \alpha \exp(-(\|r_i - r_c\|^2)/2\sigma^2)$$

ここで α は学習係数であり、 h_{ci} は近傍関数である。 r_i, r_c はそれぞれノード i, c の位置ベクトルを現わす。

以上の手順により 2 次元平面上に射影されたユーザの興味関心マップを生成する。こうして得られたユーザの興味関心マップの例が図 2 である。

図 2 の興味関心マップはあるユーザの 2005 年 9 月 1 日の Web 閲覧行動から得たものであり、特徴ベクトルは約 1600 の単語数、12 の閲覧サイト数から構成される。そして興味関心マップは 12x8 の計 96 の出力ノードとし、1 つのノードに 2 つの単語までを割当てると設定したため、特徴ベクトルの約 10% が興味関心マップに表わていることになる。

このマップには 2005 年に開催された万博愛・地球博に関連するクラスタ (図 3 左) やバスや JR といった移動手段に関連するクラスタ (図 3 右) が現れており、このユーザが万博愛・地球博に興味があったことが推測できる。

実際このユーザはこの興味関心マップを抽出した一日の間、「名古屋観光」、「名古屋駅 アクセス」、「万博 EXPO シャトル」といった検索キーワードを投入しており、それに関連した「愛 地球」、「ホテル 愛知」、「JR きっぷ」などの単語がマップ上に出現している。

³<http://mecab.sourceforge.jp/>

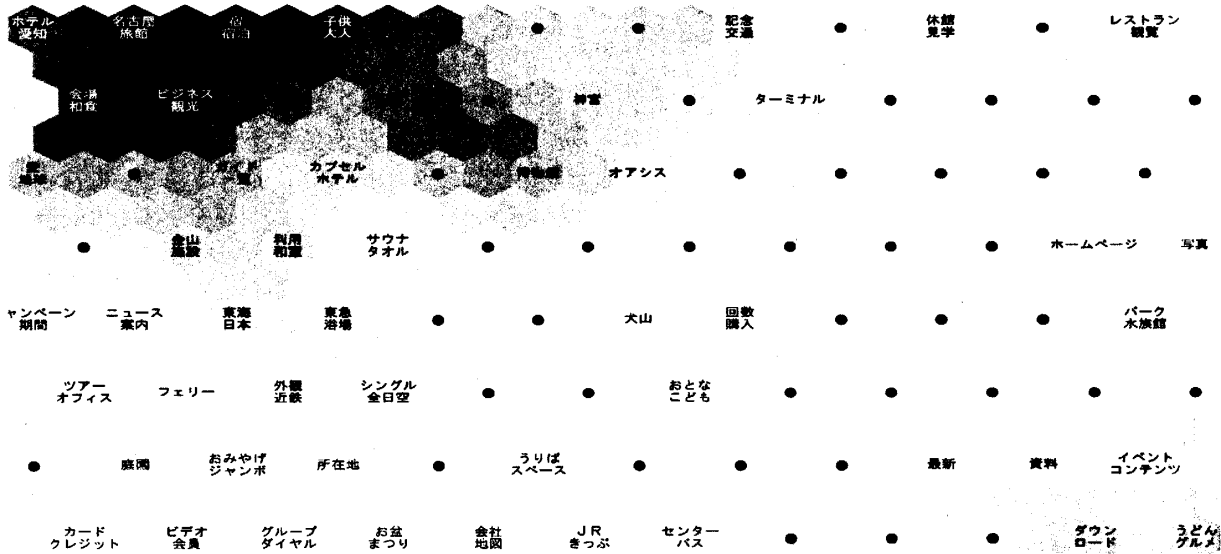


図 2: 得られたユーザの興味関心マップ

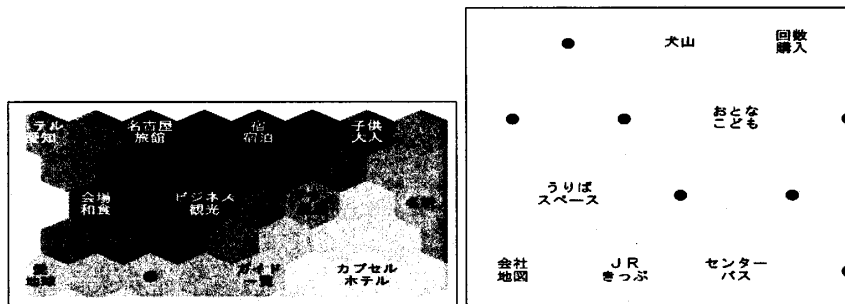


図 3: 興味関心マップ上のクラスタ

4 評価実験

ここでは得られた興味関心マップがどの程度ユーザの興味を現しているのかを評価する方法とキャンパスネットワーク内における実験結果について述べる。

4.1 実験方法

我々の目的は抽出したユーザプロファイリングがユーザの明示的な要求からどれくらいの精度を持っているかを明らかにすることである。前節でも述べたように、生成した興味関心マップはそのときのユーザの Web 閲覧の特徴的な単語がクラスタ化されている。さらにユーザの Web 閲覧行動は検索エンジンへ投入した検索キーワードを元に行なっているものと考えられるので、この検索キーワードに対応する興味関心マップ上の単語との類似度がユーザの明示的な要求を用いた興味関心マップとの評価値になると考えられる。

したがって、興味関心マップの具体的な評価方法は次の手順で行なう。

1. 興味関心マップ生成時⁴におけるユーザが検索エンジンに投入した検索キーワード K_i を保存する
2. 検索キーワード K_i を興味関心マップ上に当てはめ、該当するノードに割当てられた単語を W_k とする。
3. シソーラス^[9]を用いて K_i と W_k の類似度 S を求める。

ただし検索キーワード K_i に対応する興味関心マップ上のノードに単語が割当てられていない場合もあるため、類似度に加えて再現率 R も評価値とする。

$$R = \frac{\text{検索キーワードに対応する単語が存在した数}}{\text{検索キーワード総数}}$$

⁴興味関心マップの生成周期は一日とした

1. のユーザが検索エンジンに投入した検索キーワードの抽出は Web プロキシのアクセスログにおける Web 閲覧ログから抽出する。具体的には図 4 の形式の URL で現わされるものから検索キーワードのみを抽出し K_i とした。

```
http://www.google.co.jp/search?hl=ja&q=%E9%A6%99%E5%B7%9D
# Google に香川と入れた場合
```

図 4: 検索キーワードを投入した URL 例

検索エンジンの検索キーワードの表現にはアクセスログ解析ツール analog⁵ の表現を用いて Google, Yahoo!, MSN など代表的な検索エンジン 24 システムへ投入した検索キーワードを抽出した。

次に検索キーワードを興味関心マップにあてはめ、シソーラスを用いて該当ノードの単語との類似度を計算する方法について述べる。シソーラスを用いた単語 (概念) 間の類似度計算方法には一般的にカテゴリ間の距離を用いる方法、共通のカテゴリ数や根からの距離を用いる方法などがある [8]。

ここでは類似性判別能力が高いシソーラスにカテゴリ属性をベクトル化して表現し、2つの単語の類似度をベクトル間の内積によって計算する方法を用いる [7]。ベクトル化による類似度計算方法は次の手順で行なう。

1. シソーラスをカテゴリでベクトル化する。
2. 均等深度法 [12] によりベクトル次元数の圧縮を行なう。
3. 情報量の比により各カテゴリに重み付けを行なう (次式)。

$$W_b = W_a \times \frac{\log((C_b + 1)/N_c)}{\log(1/N_c)}$$

ただし、 W_b はノード b の重み、 W_a は初期重み、 N_c はカテゴリ総数、 C_b はノード b の下位カテゴリ総数である。

4. 検索キーワードと興味関心マップ上の対応する単語をシソーラスのベクトルで表し、 \vec{K}_i 、 \vec{W}_k とし次式により類似度 S を計算する。類似度 S は 0 ~ 1 の間で現される。

$$S = \frac{\vec{K}_i \cdot \vec{W}_k}{|\vec{K}_i| |\vec{W}_k|}$$

具体的に図 2 のユーザの興味関心マップ生成日に該当ユーザの検索キーワードとそれに対応する興味関心マップ上の単語を表 1 に示す。

表 1: 検索キーワードとマップ上の対応する単語

検索キーワード K_i	対応するマップ上の単語 W_k
名古屋駅	名古屋
観光	ビジネス
アクセス	周辺
万博	愛 ホテル

また検索キーワードにマップ上の対応する単語が複数ある場合は最も高い類似度 S のものを採用することとした。

4.2 実験結果

表 2 の期間、キャンパスネットワーク内のユーザから興味関心マップを生成し、評価実験を行なった。そしてその評価結果が表 3 である。評価値は全ユーザの平均である。

ユーザプロファイルの評価結果の比較として Foltz らはプロファイルをアンケートによる明示的な評価を行なっているが類似度は 0.67 であり [13]、本稿の自己組織化マップによる閲覧行動のクラスタ化はまずまずの結果となっていると言える。また彼らの結果では再現率は 0.25 になっているが、本稿の手法は閲覧行動が検索キーワードを元に行なわれているために彼らの結果に比べて高い値になっていると思われる。

⁵<http://www.analog.cx/>

表 2: 実験期間, ユーザ数

実験期間:	2005年8~9月
興味関心マップ抽出タイミング:	1日単位
総ユーザ数:	2,719人
ユーザ当たりの平均閲覧数(PV):	314

表 3: 評価実験結果

総検索キーワード数:	102,034
ユーザ当たりの平均検索キーワード数:	11.02
類似度 S:	0.77
再現率 R:	0.81

5 おわりに

本稿ではアクセスログをもとにユーザの大域的な Web 閲覧行動から自己組織化マップを用いて興味関心マップを抽出した。その結果見た目にも分かりやすく、ユーザの興味関心とも近い興味パターンマップが取り出せることを示した。そしてこの興味関心マップの評価にユーザが検索エンジンに投入したキーワードを用いて評価する方法を提案し、実験を行なった。

今後ユーザプロファイルのさらなる精度向上のためにユーザの持つ興味関心の変化や興味関心を長期的興味と短期的興味に対応させるなどの工夫が必要である。これはユーザプロファイルの特徴に活性度という概念を設け、活性度は閲覧したページにより刺激を受け、時間の経過と共に減衰する手法^[15]や長期の興味関心と一日限りの興味関心を分けてプロファイルを作成する手法^[14]などを適用していくことで対応できると考えられる。また抽出したユーザの興味関心パターンの遷移や他者との違いの表現方法をどのようにするのかといった問題が残されている。また本稿では一日単位でユーザの興味関心マップを生成したが本稿の興味関心マップは1ユーザあたり生成に30分弱の時間を要しておりユーザプロファイルを他のシステムと連携した応用を考える際にはユーザプロファイル構築に要する時間も問題である。

参考文献

- [1] 福原知宏; 村山敏泰; 中川裕志; 西田豊明: ウェブログ記事を用いた関心解析システム, 人工知能学会 第19回全国大会, 2C2-04, 2005.
- [2] Nanno, T., Suzuki, Y., Fujiki, T., and Okumura, M.: Automatic collection and monitoring of Japanese Weblogs., *WWW 2004 Workshop on the Weblogging Ecosystem*, 2004.
- [3] 谷口 智哉; 松尾 豊; 石塚 満: Blog コミュニティの抽出と分析, 人工知能学会, 第6回セマンティックウェブとオンтоロジー研究会, SIG-SWO-A401-08, 2004.
- [4] Kohonen, T.: *Self-Organizing Maps, 3rd Edition*, Springer-Verlag, 2001.
- [5] 橋高博行; 佐藤直之; 鈴木英明; 曾根岡昭直: パーソナライズ情報提供方式の提案と評価, 情報処理学会論文誌, Vol.40, No.1, pp.175-187, 1999.
- [6] 土方嘉徳: 情報推薦・情報フィルタリングのためのユーザプロファイリング技術, 人工知能学会誌, Vol.19, No.3, pp.365-372, 2004.
- [7] 川島 貴広; 石川 勉: 言葉の意味の類似性判別能力に関するシソーラスと概念ベースの性能比較, 人工知能学会全国大会 2D2-10, 2004.
- [8] 川島 貴広; 石川 勉: 言葉の意味に関する類似性判別能力における概念ベースとシソーラスとの性能比較, 情報処理学会 第65回全国大会, 2M-1, pp. 135-136, 2004.
- [9] 池原 悟, 他: 日本語語彙体系, 岩波書店, 1997.
- [10] 土方 嘉徳; 青木 義則; 古井 陽之助; 中島 周: マウス挙動に基づくテキスト部分抽出方式と抽出キーワードの有効性に関する検証, 情報処理学会論文誌, Vol.43, No.2, pp. 566-576, 2002.
- [11] 大野 健彦: IMPACT: 視線情報の再利用に基づくブラウジング支援法, in *Proc. of the 8th Workshop on Interactive Systems and Software (WISS'2000)*, pp.137-146, 2000.
- [12] 平川秀樹; 木村和広: 概念体系を用いた概念抽象化手法と語義判定におけるその有効性の評価, 情報処理学会論文誌, Vol.44, No.2, pp.421-432, 2003.
- [13] Foltz, P.W., Dumais, S.T.: Personalized information Delivery: An Analysis on Information Filtering Methods, *Comm. of the ACM*, Vol. 35, No. 12, pp.51-60, 1992.
- [14] 杉山一成; 波多野賢治; 吉川正俊; 植村俊亮: ユーザからの負担なく構築したプロファイルに基づく適応的 Web 情報検索, 電子情報通信学会論文誌, Vol.J87-D1 No.11, 2004.
- [15] 宮原一弘; 岡本敏雄: Web ブラウジングに基づいた興味の定量的同定法とその協調フィルタリングへの適用, 信学技法, ET97-115, pp.17-24, 1998.

Web上における『法政経済社会論文総覧』の検索閲覧システムの構築

○古隅 弘樹* 周防 節雄*

The Construction of Web-version Information Retrieval System
on the Database of "Housei Keizai Shakai Ronbun Souran"

OHiroki FURUZUMI* Setsuo SUOH*

"Housei Keizai Shakai Ronbun Souran"(Catalogue of Academic Papers on Law, Politics, Economics and Sociology) and its additional version edited by Keitaro Amano were published in 1927 and in 1928 respectively. They cover collection of indexes of important academic papers and articles on law, politics, economics and sociology published in academic journals and books during the period from the Meiji Restoration until Showa 2nd year, and they still remain one of main, important sources of information for surveying old documents of those days. We have been working on the construction of web-version information retrieval system on the database of the two indexes, the whole text information of which has been converted into XML style. Our system is now released on our website. In the present paper we describe how our system has been developed and how it can be used.

1. はじめに

昭和2年発行の天野敬太郎編『法政経済社会論文総覧』は明治・大正・昭和初期の法律・政治・経済・社会に関する専門雑誌や、大学紀要・論集に掲載された論文記事目録の集成であり、五十音順配列の見出し語(テーマ)によって分類されている。ここには明治初期～大正15年6月の雑誌論文・記事が収録されている。更に、昭和3年には、大正15年7月～昭和2年の論文記事目録を新たに収録した「追篇」が発行された(以下ではそれぞれを「天野正篇」、「天野追篇」と呼ぶ)。大部分の関係資料がその創刊号から収録されているため、今でも明治・大正・昭和初期の文献調査のための重要資料となっている。

著者らが主宰する書誌情報データベース研究会(代表者:周防節雄)では、「日本社会科学形成期書誌索引情報集成データベース」を構築するプロジェクトを推進中で、手始めに、天野正篇・追篇の全文入力データを作成し、この入力データを計算機処理しやすいようにXML文書化した。このXML文書を用いてデータベースを構築して、Web技術を連携させることで天野正篇・追篇に収録された論文記事目録の検索閲覧システムの開発を行った。このシステムは現在も開発を継続中であるが、機能の一部を一般公開している¹。本稿では、公開中のシステムの開発手法や検索機能について述べる。

2. データ入力とXML文書化

天野正篇・追篇の本文の情報をプログラム処理するために、データの入力規則を指定して、入力業者に委託して作成したものが初期入力ファイルである。天野正篇本文の冒頭ページを図1に、これに対応する部分の初期入力ファイルの内容を図2に示す。

*兵庫県立大学 神戸学園都市学術情報館

Academic Information Systems Centre, Kobe Gakuentoshi Campus, University of Hyogo

¹ (公開URL) <http://mighty.gk.u-hyogo.ac.jp/dabii-jass/>

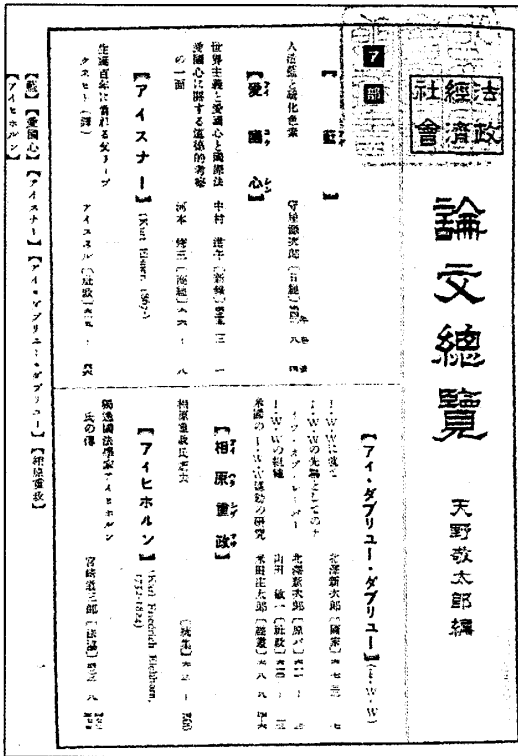


図1 天野正篤本文の冒頭ページ

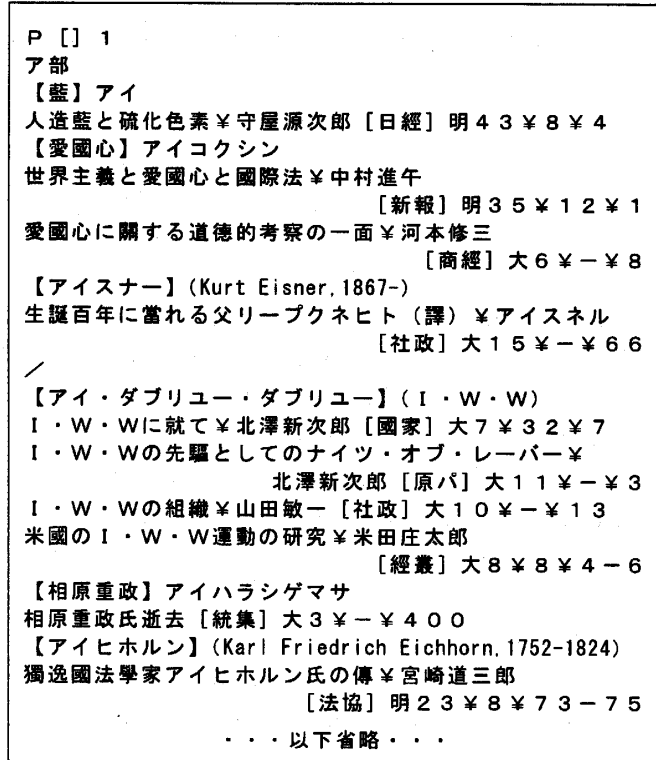


図2 初期入力ファイルの冒頭部分

初期入力ファイルでは、可能な限り正字（原本で使用されている漢字）で入力している。原本では書誌情報の記載レイアウトが統一されており、①タイトル、②執筆者名、③掲載雑誌・論集名（略語）、④発表年（和暦）、⑤巻、⑥号、の順に記述されている。プログラム処理に必要な情報の区切り記号（「¥」など）を挿入した他は、原本の記載どおりに入力した。ファイルの文字エンコーディングはShift_JISを使用した²が、可能な限り原本の記載と同じ漢字を使用するため、最終的なアプリケーションやデータベースではUnicodeの文字セットを使用することを想定し、Shift_JISでは入力不可能であってもUnicodeでサポートされる文字については、HTMLにおける文字実体参照の記述形式を用いて文字コードを埋め込んだ。

本文の記述構造に基づいてXML文書構造を独自に設計し、初期入力ファイルをプログラム処理してXML化を行った（図3）。Unicodeの文字セットを扱えるようXML文書におけるエンコーディングはUTF-8を使用した。

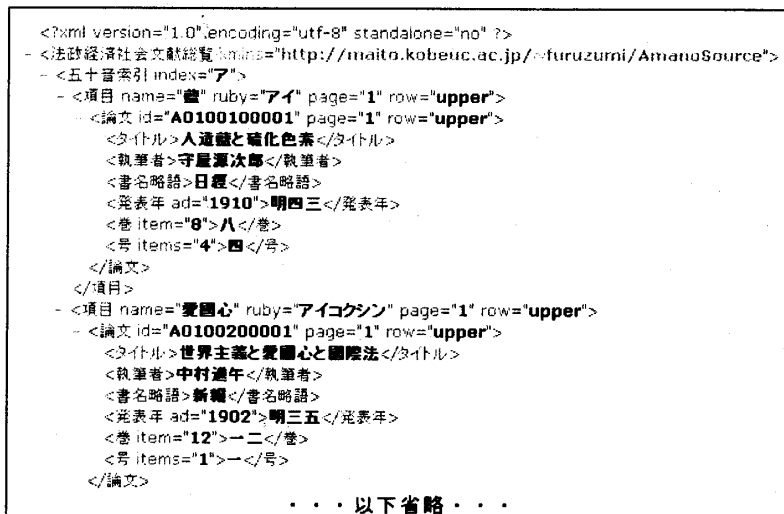


図3 XMLファイルの冒頭部分

² Unicodeのエンコーディングに不完全な対応のテキストエディタを使用すると、データが破損する恐れがあるため、入力データの校正作業におけるトラブルを避けるためにShift_JISを使用した。

4. XML データベース

データベースの実装にはXMLデータベース (eXist³) を使用しており、天野正篇および追篇本文のXMLファイルをそのまま格納している。初期入力ファイルは正字で入力されているため、本文のXMLファイルも正字であるが、異体字等の漢字を一つの字に統一することによって表記揺れを減らし、可能な限り新字に置換したXMLファイルを用意して二重化することで、検索時における問題点の解決を図っている。コレクション階層及びドキュメントの配置はデータベースの検索効率を考えて、新字 (simplified) と正字 (traditional) で分けてXML文書を管理している⁴ (図4)。

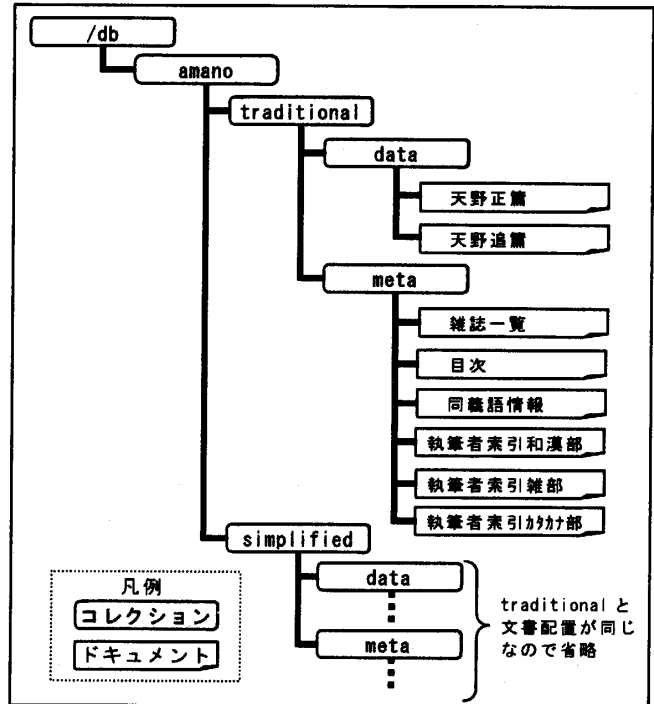


図4 コレクション階層とドキュメントの配置

5. Web 検索閲覧システム

データベースの公開ページ (図5、図6) から本システムを利用できる。本システムは天野正篇・追篇の内容を格納したXMLデータベースと、その内容を検索・表示するプログラムで構成され、これらは共にJavaサーブレットとして動作するWebアプリケーションとして構築した。以下では、公開している検索機能について紹介する。

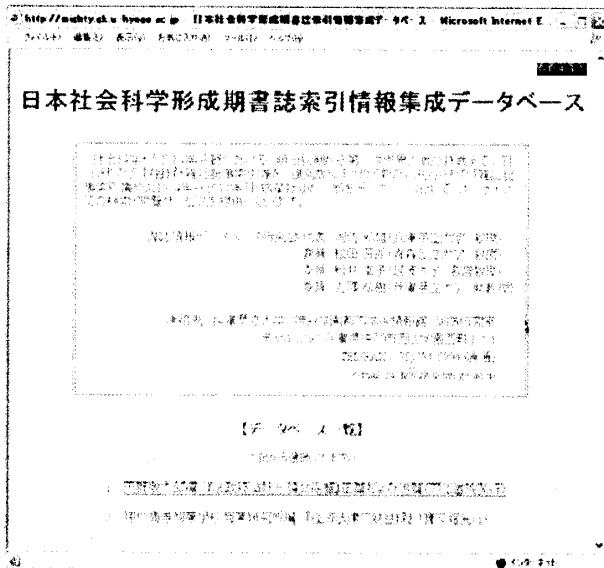


図5 データベース公開ページ

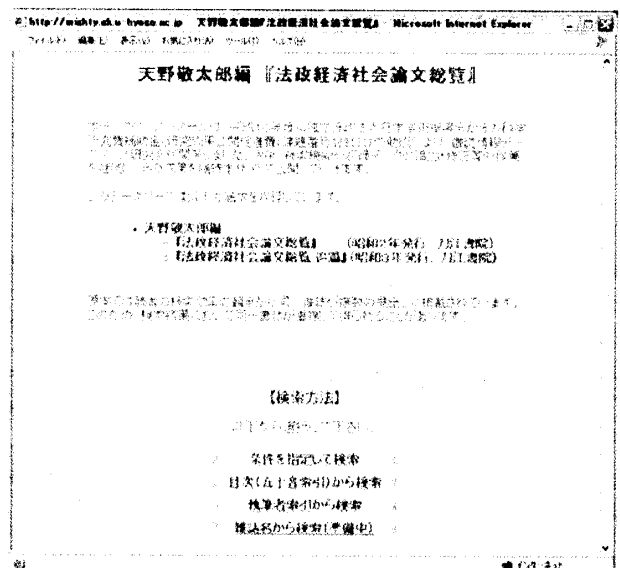


図6 「法政経済社会論文総覧」検索ページ

³ eXistは、XPathやXQueryによるクエリサービスをサポートするオープン・ソースのネイティブXMLデータベースである。現在のシステムでは2006年1月24日付けのスナップショットを使用している。詳細は<http://exist.sourceforge.net/> を参照。

⁴ eXistはXML文書をファイルシステムに似た機構で管理している。フォルダに相当するものを「コレクション」、XMLファイルを「ドキュメント」と呼び、XML:DB仕様に基づいている。

5.1 条件を指定して検索

データベースの内容に対して条件を指定して検索することができる。指定できる項目は図7に示すとおりで、発表年以外の項目については指定文字列の部分一致で検索し、各条件はAND結合になっている。

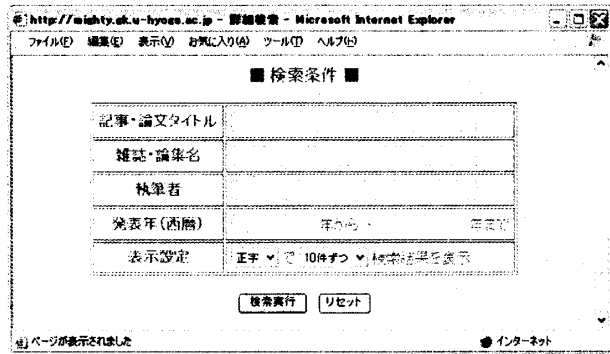


図7 「条件を指定して検索」入力項目

5.2 目次から検索

天野正篇・追篇では、論文記事目録はその内容に応じて見出し項目(テーマ)毎に分類されている。見出し項目はヨミの五十音順に配列されており、見出し項目の一覧、すなわち目次を表示することで、原本の閲覧と同様の検索機能を提供している。Webページで用意した五十音索引表(図8)から、ヨミ毎に見出し項目の一覧(図9)を表示することができ、選択した見出し項目から論文・記事を検索できる。また、本文中に記載されていた見出し項目の関連情報を利用して、他の関連見出しへのリンク等の情報も併せて表示している。

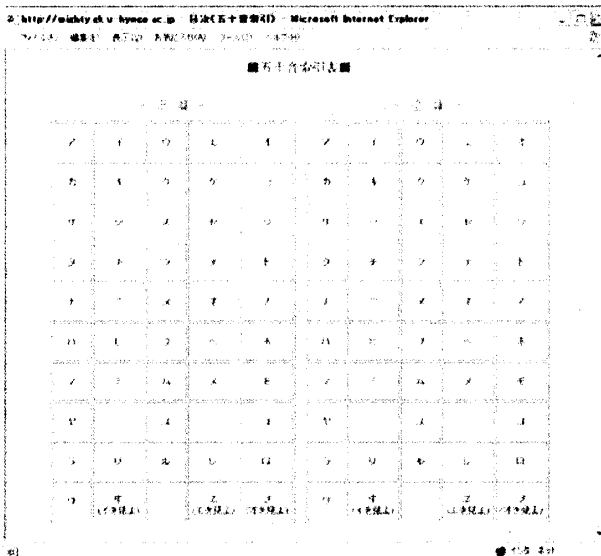


図8 「目次から検索」五十音索引表

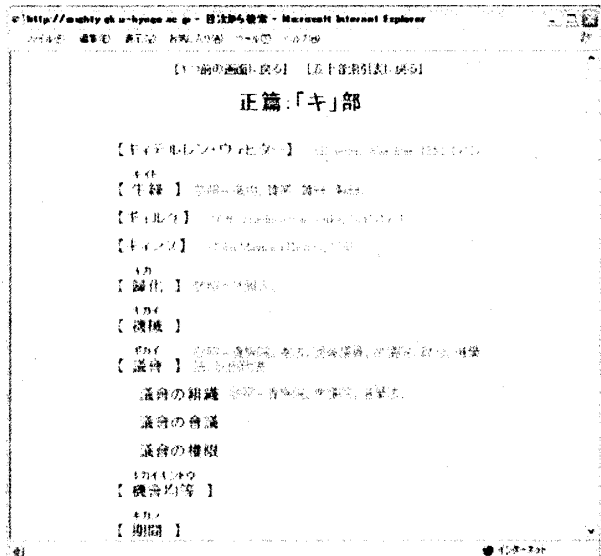


図9 見出し一覧の表示例(正篇:「キ」部)

5.3 執筆者索引から検索

天野追篇の巻末に正篇・追篇を通じた執筆者索引が掲載されている。執筆者はヨミの五十音順に配列されており、執筆者毎に論文記事目録が記載された正篇および追篇本文のページ番号を調べることができる。ただし、漢字表記の氏名だけが五十音順に配列されており、英字や記号で始まる執筆者は「雑部」に分類されている。なお、カタカナで始まる氏名の執筆者は原本の執筆者索引には掲載されていなかったため、新たに「カタカナ部」として加えた。執筆者索引の入力データと、本文のXMLデータの情報を照合したところ、索引氏名の不整合や索引漏れが多数見つかったため、これらを修正、整理する作業が必要となった。この検索機能では索引表(図10)で選択したヨミで始まる執筆者の一覧表示(図11)から、特定の執筆者に関する論文記事を検索することができる。

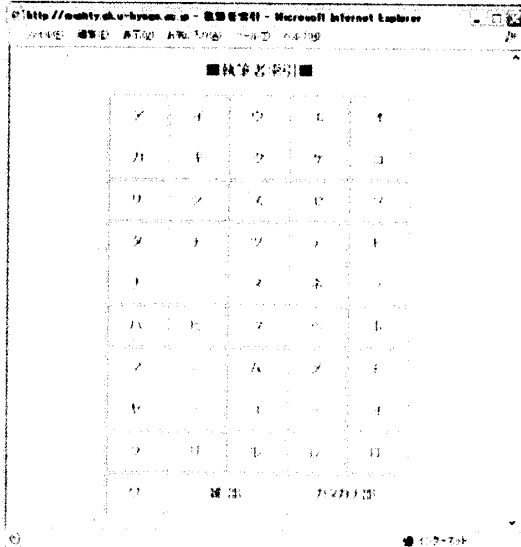


図10 「執筆者索引から検索」索引表

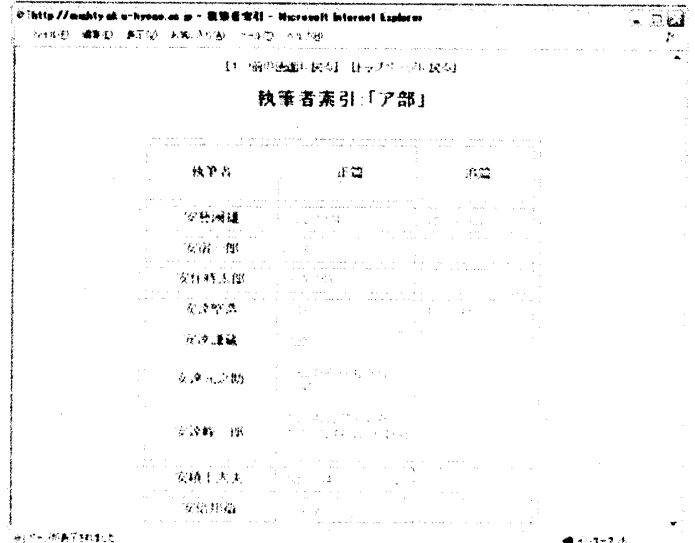


図11 執筆者索引の表示例（ア部）

5.4 検索結果の表示

表示に使用する漢字(正字または新字)、および検索結果の一覧表示の一頁あたりの件数(10件または20件)を各検索ページで指定することができる。検索結果については、ヒットした件数と論文記事の一覧を指定件数ずつ表示する(図12)。

各レコードの「詳細表示」ボタンで更に詳細な情報を閲覧できる(図13)。論文記事によっては、同一タイトルが複数の雑誌巻号で掲載(あるいは連載)されたことを表現する複雑な構造を持つ情報もある(図14)。

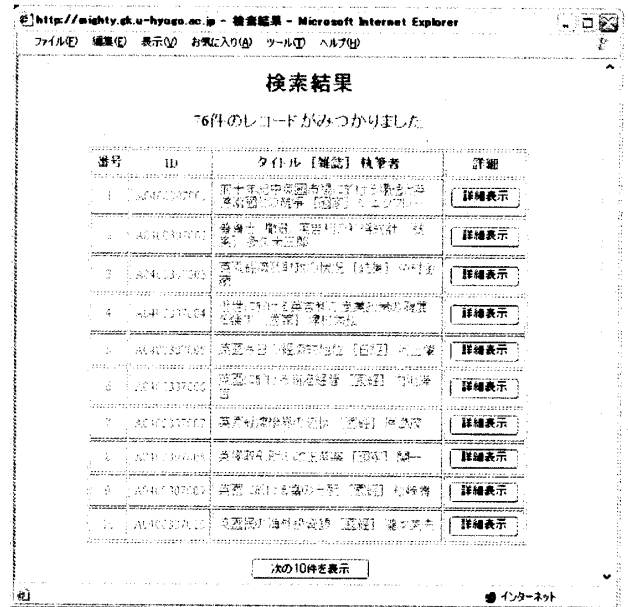


図12 検索結果の一覧表示例

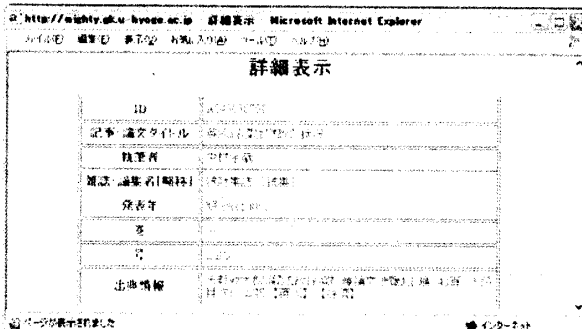


図13 詳細表示例(1)

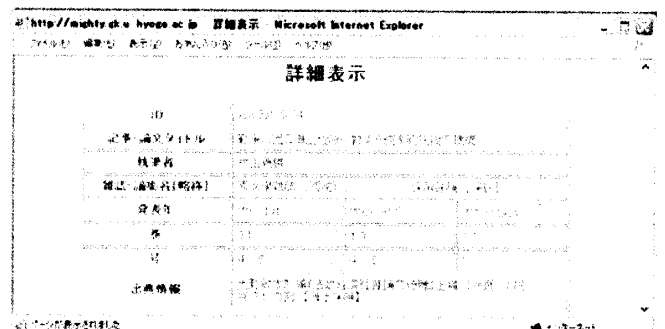


図14 詳細表示例(2)

6. 今後の予定と課題

XMLの文書構造については、天野正篇・追篇における本文の記述内容や構造を可能な限り原本どおりに表現することを主眼に置いたため、独自に定義したものを使用したが、検索効率の向上や新たな検索機能の追加、他のデータベースやアプリケーションと

の連携など、今後のシステム開発に向けた記述形式の見直しや変換規則の定義について再検討したいと考えている。

システムの検索機能については、現在公開中の機能に加えて、目録情報における掲載雑誌・論集名から検索する機能を構築するためにデータを整理中である。具体的には、目録情報にある発表年と雑誌の巻号の情報は対応するため、これらを整理すれば、雑誌毎にある程度の記事索引を作成することができる。この他にも、各検索機能間の連携や検索結果の絞り込み機能など、追加したい機能は多い。

データベースで収録している目録情報には、執筆者の氏名が索引とは異なることや、掲載雑誌の巻号と発表年の対応が一致しないなど、明らかに記述内容の修正が必要なものがこれまでの作業で多く見つかっている。また、天野正篇・追篇では読者の検索効率の観点から同一書誌が複数の見出しに掲載されているので、検索した結果、同一書誌が重複して出てくることがあり、冗長な情報表示を排除できるように対処する必要がある。

現在、神戸高等商業学校商業研究所編『経済法律文献目録』（寶文館発行）のデータベース化作業も平行して進めているが、目録情報の収録時期が天野正篇・追篇と重なるので、収録内容の照合作業に利用でき、重複書誌をコントロールすることによって統合データベースを構築することも構想している。

本システムは未だ開発途上であり、新たに追加する検索機能やデータベースの収録内容の更新分などを逐次公開していく予定である。是非現行システムをご利用頂き、システムに関する要望やデータに関する不具合などをご連絡いただければ幸いである。

謝辞

本研究は、(独)日本学術振興会の平成16年度科学研究費補助金（研究成果公開促進費：課題番号168131：作成代表者[周防節雄]）により行われた。書誌情報データベース研究会のメンバーである松田芳郎青森公立大学教授（一橋大学名誉教授、東京国際大学名誉教授）から底本の選定やデータベース構築の基本方針に関する貴重な助言を頂いた。同メンバーの松井幸子筑波大学名誉教授からは、原本目録情報の記載形式の調査などで多くのご助力を頂いた。また、天野正篇・追篇の復刻版出版元である皓星社の藤巻修一氏からも助言を頂いた。ここに記して謝意に替いたい。

参考文献

- [1] 天野敬太郎編『法政経済社会論文総覧』、刀江書院、1927年
- [2] 天野敬太郎編『法政経済社会論文総覧 追篇』、刀江書院、1928年
- [3] 石山洋他編『明治・大正・昭和前期 雑誌記事索引集成』、社会科学編 第7～9巻、皓星社、1994～1997年（天野正篇・追篇の復刻版）
- [4] 周防節雄・古隅弘樹、『「統計資料解題」のデジタル化に関する技法とシステム構築』、神戸商科大学研究叢書LXIX、神戸商科大学経済研究所、2003年。
- [5] 周防節雄・古隅弘樹、『旧植民地に関する統計書誌情報データベースシステムの構築』、兵庫県立大学経済経営研究叢書LXVII、兵庫県立大学経済経営研究所、2006年。
- [6] クン・イー・ファン著、原隆文訳、『XSLT Web 開発ガイド』、ピアソン・エデュケーション2001年。
- [7] 屋内恭輔・安倍隆明、『XML スキーマ書法』、毎日コミュニケーションズ、2003年
- [8] Dan Chang・Dan Harkey 著、高際理監修、石井史子訳、『Java & XML データアクセスガイド』、翔泳社、2000年。

科学技術文献検索システムにおける大規模用語辞書の活用について

○甲田彰¹, 森田歌子²

¹ 科学技術振興機構 情報提供部

² 科学技術振興機構 審議役

We, Japan Science and Technology Agency (JST) started to provide JDream II since April 1st, 2006. It is a new document retrieval system that is integrated with JOIS that JST had provided since 1976. This system has a large-scale scientific and technological dictionary to support users' advanced search. In this report, I describe the way to support advanced search by this dictionary and the future enhance, too.

1. はじめに

科学技術振興機構（以下「JST」）は2006年4月から、新たな科学技術文献検索サービス「JDream II」の提供を開始した。これは、1976年にサービスを開始したJOISと、2003年に主に大学や病院等の利用者向けに開発しサービスを提供してきたJDreamを統合し、機能を飛躍的に向上させた新しいシステムである。

JDream IIは、従来の検索システムの機能を損なうことなく、利用者の要望に応じて様々な検索支援機能を搭載した。そのうちの一つが今回報告する大規模辞書およびそれを用いたJSTシソーラスブラウザである（詳細後述）。

JSTは、国内外の科学技術文献と国内の医学関連文献を収集し、概要（抄録）の作成、索引語の付与等の加工を行っており、索引語を用いて検索することにより、漏れのない精度の高い検索を行うことが可能である。

そこで、JDream IIでは、JSTが付与した索引語をより容易に利用できるよう、大規模用語辞書を活用した検索支援機能「JSTシソーラスブラウザ」を開発した。本稿では、JSTシソーラスブラウザによる大規模用語辞書の活用について、その目的やデータ構成、システム構成を報告する。

2. JSTシソーラスブラウザの目的

従前の検索システムでは、索引語の参照機能は用意されていたが、利用者があらかじめ索引語に見当をつけて参照する方法を採っており、JSTの付与する索引語を全く知らない利用者にとって、索引語の利用は必ずしも容易ではなかった。

これに対して、JDream IIでは、利用者がより簡便に索引語を使用できるよう、検索支援機能「JSTシソーラスブラウザ」を用意した。JSTシソーラスブラウザは、利用者到大規模用語辞書を提供するためのインタフェースであり、漏れなくかつ高い精度の検索の実現を目的としている。

3. JSTシソーラスブラウザの機能

3.1 データ構成

3.1.1 大規模用語辞書の収録内容

大規模用語辞書の2006年4月時点での収録用語数は約40万語であり、収録内容は以下の通りである。

- (1) シソーラス等の用語辞書および最近5年間に出現した索引頻度30以上の準シソーラス用語とその同義語、異表記語
- (2) 自然語とシソーラス用語との包摂関係も、一部収録している。

また、今後の方針としては、索引付けでの出現頻度に関わりなく、索引で用いられた用語は漏れなく収録する方針であるが、当面は索引(すなわち文献)上に一定の回数以上出現する重要用語を収録する方針である。

3.1.2 索引側と検索側の共通の辞書

大規模用語辞書は、単に収録用語数が多いだけではなく、索引側と検索側が同じ辞書を共有するところに特長がある。

フーグマンの5命題理論を検索サイドから見直した「自明な5原則ルール」では、漏れを減らすための第4原則ルールを「検索者が検索式を作る際にDB仕様(インデキシングの仕方)を容易に思いつくようになるほど、再現率が高くなる」と解釈している[1]。

図2の通り、JDream IIの利用者は、大規模用語辞書を用いることで「インデキシングの仕方」に沿って検索することが可能である。

3.1.3 大規模用語辞書のメンテナンス

索引側は、抄録やタイトルから、重要と思われる単語を抽出し、大規模用語辞書を用いて索引語を付与するが、大規模用語辞書にない新語が見つかった場合は辞書メンテナンスシステムを用いて辞書の用語を補強していく予定である。

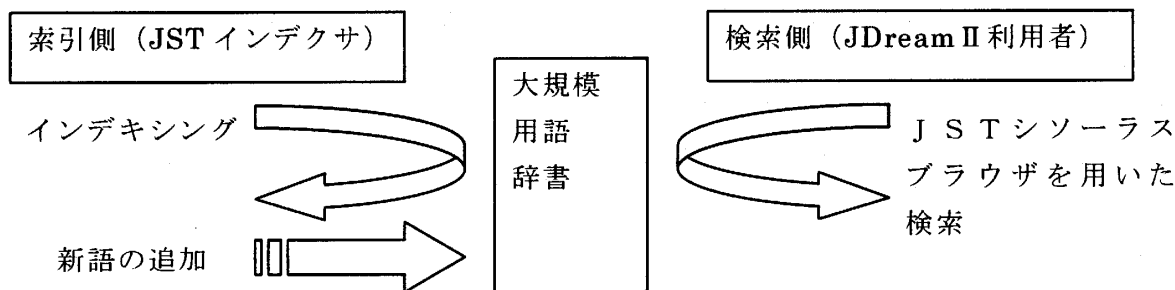


図2. 大規模用語辞書

3.2 システム構成

JST シソーラスブラウザは以下の機能から構成されている。

(1) 候補語一覧の表示

検索システムの利用者が入力した単語に対して、大規模用語辞書に搭載された科学技術用語と、それらに関連する索引語を一覧で表示する。索引語を表示する際は、索引語の種類(シソーラス用語、順シソーラス用語、化学物質名)も表示する。

大規模用語辞書を検索するにあたっては、「前方一致」「完全一致」「部分一致」のいずれかが選択可能である。図3は「抗がん」という単語から「前方一致」で大規模用語辞書を検索した結果の一覧画面の例である。

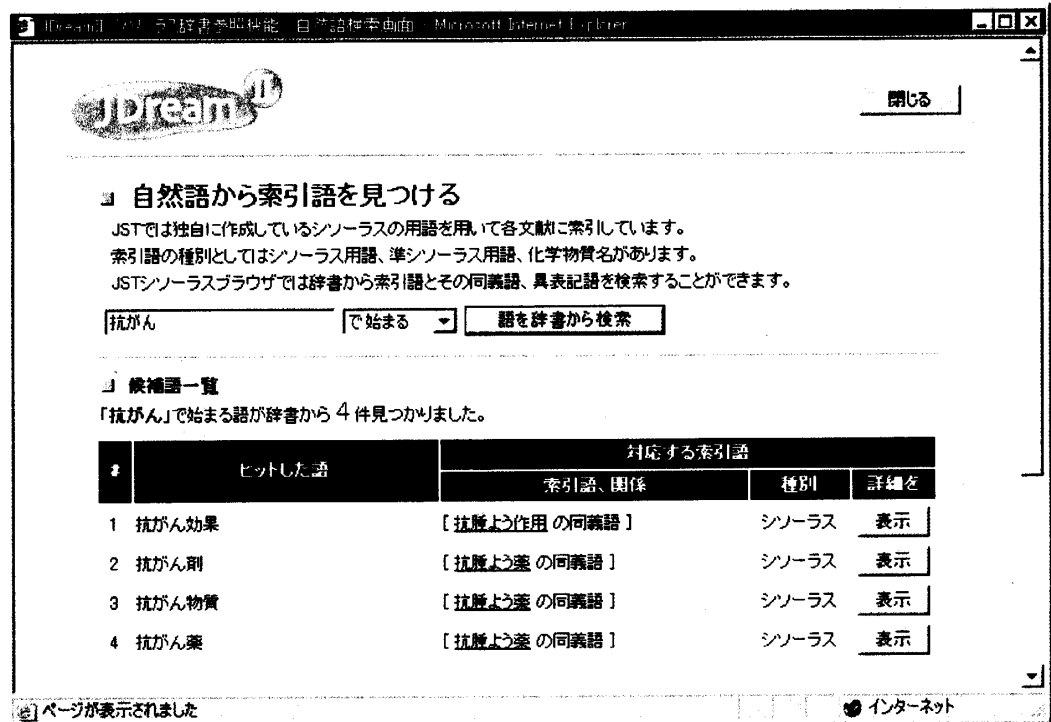


図 3. 一覧表示画面の例

(2) 選択した索引語に関する同義語等の表示

図 3 の一覧表示画面には、ヒットした大規模用語辞書の用語とそれぞれに対する索引語が表示されている。この中から索引語を選べると、その同義語を表示する。図 4-1 はシソーラス用語「抗腫瘍薬」を選択し、その同義語を表示した画面の例である。

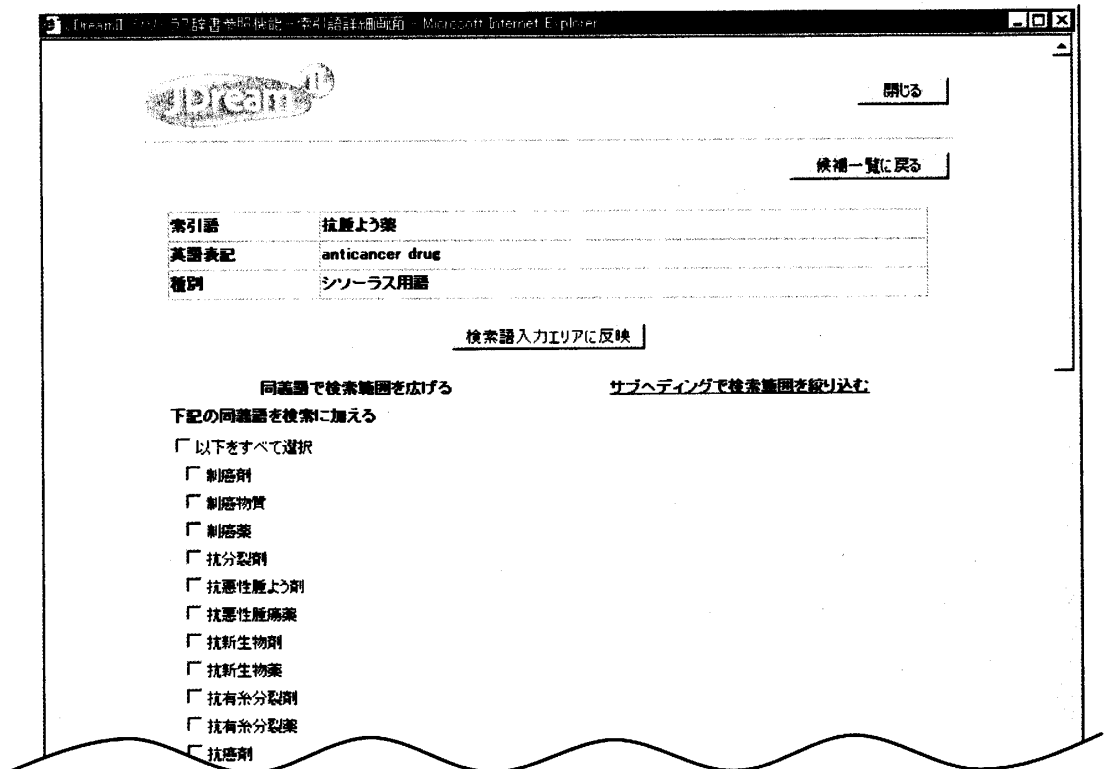


図 4-1. 索引語詳細画面（同義語の表示）例

また、同義語表示画面（図 4-1）の下には、索引語がシソーラス用語の場合は上位語、

下位語、関連語を表示し、準シソーラス用語や化学物質名の場合は上位語、関連語を表示する。

さらに、表示された上位語、下位語、関連語を選択すると、当該用語を中心とした包摂関係を表示する。図4-2はシソーラス用語「抗腫瘍薬」の直近上位語、直近下位語、関連語を表示した画面の例である。

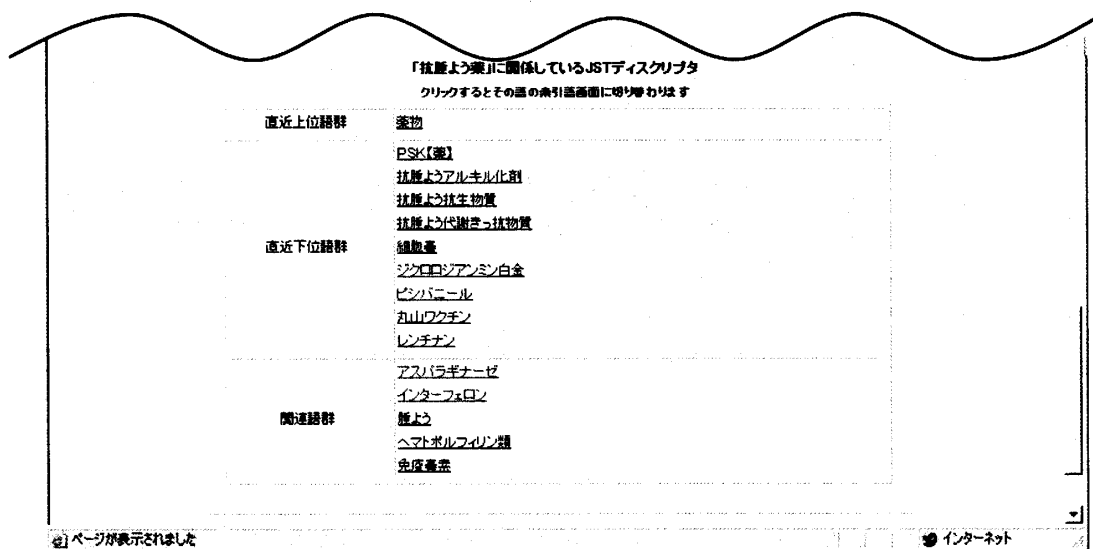


図4-2. 索引語詳細画面（包摂関係の表示）例

4. 今後の課題

4.1 シソーラス閲覧機能の改善

現在のJSTシソーラスブラウザは、索引語の階層関係や関連語の関係を一覧で表示する形を採っていないが、JDream IIの利用者からは、「用語間の相互の関係を俯瞰したい」という要望が寄せられている。今後の課題として、検討していきたい。

4.2 検索行動に基づく大規模用語辞書の充実

図2の通り、大規模用語辞書の用語の拡充は索引作業の中で出現した新語を追加して行う予定である。ただし、利用者側が検索に使用した単語も、辞書を充実するための重要な要素である。JDream IIのサービス開始後一定期間が経過した段階で、検索行動を分析し、大規模用語辞書に単語を追加していきたい。

参考文献

[1] 桐山勉, 長谷川正好, 川島順, 玉置研一, 吉田郁夫, 辻川剛由. インデキシングと検索の両視点から特許情報検索システムの今後の方向を探る(考察と提案), 情報の科学と技術, vol.53, no.3, 2003, p.152-158

ミュージアム資料情報構造化モデル応用の検討

Examination of application of Structured Model for Museum Object Information

齋藤伸雄[†]
Nobuo Saito

博物館や美術館などにおける業務支援と情報共有を目標として、東京国立博物館が中心となり、ミュージアム資料情報構造化モデルが策定された。そこで、このモデルを応用して収蔵品管理システムを試作し、モデルのシステム適合性について検討を行った。本稿ではミュージアム資料情報構造化モデルについて紹介し、システムへの応用に関する検討について報告する。

"Structured Model for Museum Object Information" was settled on by the group that centered on Tokyo National Museum aiming at sharing information between museums and work support. Then, the goods management system was made for trial purposes by applying this model, and the system adaptability of the model was examined. It introduces "Structured Model for Museum Object Information", and it reports on the examination concerning this model's application to the system.

1. はじめに

博物館や美術館・資料館など(以下「ミュージアム」)において取扱う対象(以下「資料」)は、多種多様である。そしてミュージアムの規模にもよるが、多い所では10万点を超える資料を保有しており、ミュージアムはそれらの資料をさまざまな目的のために維持管理する責務を負っている。多種多様な資料を管理する情報項目については、ミュージアム間の資料情報の共有化や様々な対象への情報公開のために標準的な指標に基いて管理されることが望ましい。そこで、東京国立博物館の博物館情報処理に関する調査研究プロジェクトチームによって、「ミュージアム資料情報構造化モデル」(以下「モデル」)が策定され、平成17年11月に提案された。モデルの仕様書[1]によれば、このモデルはミュージアムにおいて資料情報システムを開発する基盤として利用されることが目的とされており、さらに特定の実装には依存しないことを意識して策定されている。即ちこのモデルを応用して様々な実装が行われることが期待されている。

ところが、モデルでは、資料に関する多様な管理対象情報項目を抽象化して策定されており、実装されたシステムによって提供されるサービスにおいて、モデルに盛り込まれた情報項目が情報基盤として必要十分な記述能力を有しているかどうか評価される事が必要であると考えられる。そこで、モデルに基づいた仕様で、ミュージアムにて活用される事を想定した収蔵品管理システムを試作し、適合性の検討を行うこととした。

[†]凸版印刷株式会社 総合研究所
Corporate R&D Division, TOPPAN Printing Co., Ltd.

2. ミュージアムにおける情報管理の背景

ミュージアムにおいて所蔵資料の状況を把握するという、当然と思われる事が、実はミュージアムにとって最も難しいことであると言われている[2]。ミュージアムでは、多種多様な資料を所蔵し、業務の中で発生するさまざまな要件に合わせて必要な情報を迅速に提供することは重要な課題である。ところが、村田の報告[3]によると、ミュージアムにおける資料情報の公開や共有については、これまでに様々な試みが行われたが、現在までのところ実用的な情報共有は実現されておらず、webによる情報公開も一部のミュージアムが実施しているに留まっている。そこには、各ミュージアムが独自に管理対象情報項目を検討し、情報システムを開発あるいは導入しているという事情が存在し、共通の指標を用いた情報の共有が実現されづらい状況にある。

また、これまでにミュージアムによって参照可能な標準が存在していなかった訳ではなく、国際的な博物館関連組織や、国別の博物館関連協会などにより標準が提案されている(表1)。しかし、それらが国内のミュージアムによって普及していない背景には、ミュージアムにおいて実際に運用されている資料情報や情報システムのあり方と、提案された標準との間に多くの細かいギャップが存在していると言われている[4]。

表1 主な博物館情報標準

名称	提唱団体	概要
博物館資料情報のための国際ガイドライン： CIDOC 情報カテゴリ	ICOM/CIDOC	ミュージアム資料を記述するための国際標準。受入から廃棄に至る業務プロセスを22の情報グループを74の情報カテゴリに分類。
CIDOC/CRM	ICOM/CIDOC	情報カテゴリを盛り込んだ概念参照モデル。オントロジを用いて各データの関係性を記述する。RDF/XMLで記述される。
MDA SPECTRUM	MDA	ミュージアムにおける収蔵品管理手順を20の項目で記述する。
Dublin Core	DCMI	情報資源発見のためのメタデータ記述規則。15の基本エレメントセットから拡張可能。

3. モデルの概要

3.1. 対象

既に述べた背景を踏まえ、このモデルは、ミュージアムによって継続的にデータがメンテナンスされる事が不可欠であるとの認識のもと[5]、ミュージアム関係者を中心としたプロジェクトチームによって検討された。ミュージアムによって一貫した資料情報の管理を行い、それによってミュージアムにおける様々な業務を支援し、資料情報を広く公開し、複数のミュージアムによる資料情報の共有・横断的な利活用を円滑に行うといった多角的な活用を視野に入れて策定されている。しかし対象としている資料は、歴史・民族・考古・美術等の、移動可能な人工物とされており、鉱物・生物標本といった自然科学資料や、寺院などの建造物、無形文化財は対象としていないことが明示されている。さらに、収蔵品と言われるような一次資料を対象としており、参考文献や資料を撮影した画像などの二次資料の情報は、参照に必要な情報を考慮するまでに留められている。

3.2. 属性

このモデルでは記述対象資料について34種類の「属性」が定義されており、各属性について

出現回数(必須/任意の別も含む), や詳細要素が与えられている。「属性」は, その性格によって4つに分類されている(表2). 属性によっては統制語彙の使用が推奨されるものもあるが, 統制語彙で用いる言葉をこのモデルで示すものではなく, 利用者に任せることとなっている。

表2 モデルの「属性」と性格による分類

性格	概要	属性
識別・特定	資料を識別したり分類するために管理される情報	識別子, 資料番号, 名称, 分類, 用途, 様式
物理的特性	資料の物理的な Spec を記述する情報	品質形状, 材質, 技法, 形状, 員数, 計測値, 部分, 保存状態, 付属品, 印章・銘記
履歴	資料が創造され, 整理処分されるまでの取り扱い履歴に関する情報	制作, 出土・発見, 来歴, 取得, 整理・処分, 受入, 調査, 修復, 展示, 所在, 価格評価, 受賞・指定
関連・参照	資料に関連する情報	権利, 関連資料, 文献, 画像, 記述ノート, 記述作成

3.3. 構造

ミュージアム資料には, 多くの構成物から成るものもあり, 全体を管理対象とする場合もあれば, 部分のみを展示するといった場合もある. 各属性は資料の単位毎に作成されることとなるが, 資料の単位をどのように捉えるかが問題になることがある. そこで, このモデルでは資料情報を構造的に記述することを可能とし, 不可分な資料を最小単位として, 資料の集合に対しても属性を記述することを可能とするような仕様となっている(図1). また, 資料に関する研究をすすめているうちに, 資料単位を分割する必要性が生じる場合もあり, その場合にも新たな子構造を記述することが可能となっている. 最上位の記述単位は, システムによる管理対象の基準となるが, 通常ミュージアム業務の中では, 資料の受入時に記録する単位に等しいと考えられることから, 「受入単位」と呼び, 受入単位の間には階層関係を持たないこととされている。

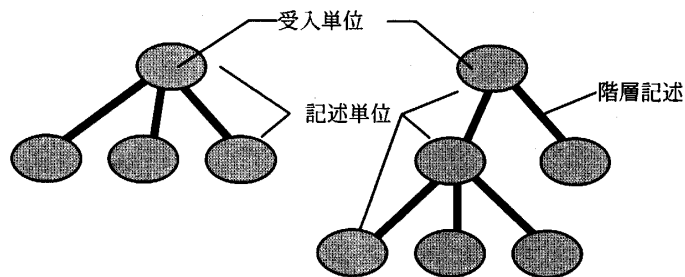


図1. 資料情報の構造的な記述

4. モデルの応用

秋元らの研究によると, 一般的な収蔵品管理システムの機能は11に分類できる[6]. 本研究でのシステム試作においては, この機能分類を参考に機能を絞り, モデルの適応性を確認するために必要な「収蔵品情報」と「貸借情報」の2つを実装した。

試作システムは, ネットワークを活用して利活用されることを想定し, web アプリケーションにより実装を行った. また, このモデルのデータスキーマをそのままデータベースに用いるのではなく, データスキーマの改良や拡張が行われる事を考慮して, モデルをクラス化し, RDB のスキーマへとマッピングして格納する方式を採った(図2). 但し, アプリケーションによって呈示されるインタフェース上ではモデルの属性名を用いて項目を表示した。

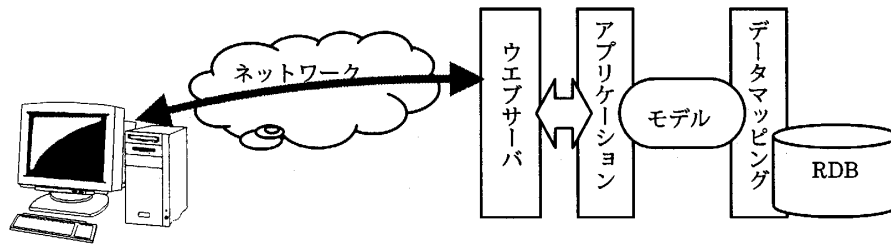


図2. 応用システム

5. システム適合性に関する考察

試作システムを試験的にミュージアム業務に適用してみると、管理項目については業務上各ミュージアム固有の項目が必要となる場合が多く、試作システムのようにモデルの属性名そのままを管理項目とした場合には不足が起きる事がわかった。さらに、識別子には分類情報を含めるなど特別な意味を持たせて管理している場合もあり、モデルのように別の属性として取扱うと却って業務が煩雑になる場合も想定される。さらに、ミュージアムによっては取扱う資料が多様であり、分類に階層を持つ場合も想定され、記述される内容に関してミュージアム間での相互互換性に憂慮すべき点があると考えられる。ただし、これらの懸案点はモデルの能力によるものではなく、その応用や活用の方法によるものである。

6. まとめ

ミュージアム資料の情報記述の標準として広く利活用されるべくミュージアム資料情報構造化モデルが策定された。本研究では、このモデルを用いた収蔵品管理システムを試作し、モデルの応用にあたっての課題の抽出を試みた。そしてモデルの応用にあたっては多様なミュージアム業務の特徴を考慮する必要があるとの知見を得た。情報化の促進やIT技術を活用した業務効率化の流れの中で、ミュージアムにおける資料の管理へ情報システムが活用されてゆくであろうと想像される。モデルを応用したミュージアムの情報化に有用な成果を導き出すべく、引き続き情報システムにおけるモデルの応用にあたっての課題の検討を深めてゆきたいと考えている。

参考文献

- [1]東京国立博物館：ミュージアム資料情報構造化モデル,
<http://webarchives.tnm.jp/download/informatics/smmoi20051216.pdf>
- [2]石森秀三：博物館資料論，初版，(財)放送大学教育振興会，pp.98-99，2000
- [3]村田良二：ミュージアム資料情報構造化モデルによる博物館業務支援と情報共有，(社)情報処理学会 研究報告 DD-53，pp.9-16，2006
- [4]村田良二：ミュージアム資料情報構造化モデルの開発，(社)情報処理学会 人文科学とコンピュータシンポジウム 論文集 平成 17 年 12 月，pp.63，2005
- [5]村田良二：ミュージアム資料情報構造化モデルによる博物館情報のためのデータベースの試作，画像電子学会 第 4 回画像ミュージアム研究会 研究報告，2006
- [6]秋元良仁 鈴木理洋：博物館情報の相互利用を目的とした文化財情報システムの提案，(社)情報処理学会 人文科学とコンピュータシンポジウム 論文集 平成 16 年 12 月，pp.99-106，2004

脳の情報処理と問題解決

Information Processing of Brain and Problem Solving

福永征夫 Masao FUKUNAGA

ABSTRACT : To recover the damaged environment and ecosystem of the earth, we should search and discover broader domain knowledge which is common among different domain knowledges, and through it, we should maintain continuity between every diverse activities of human beings to live, that is, we have to let those activities not be inconsistent with each other to economize the amount of resource and energy.

The characteristic of brain's information processing function is maintaining continual understanding and practice to the living environment by the process to realize concurrent optimization of each part and the whole between plural diverse and different domain knowledges in order to discover common broader domain knowledge to put it into practice.

1. 地球規模の難題への対処

多様な自然の破壊や自律的な人間の精神の荒廃を伴う地球規模の深刻な問題群に有効に対処するためには、人間の脳の情報処理の特質を解明することが不可欠である。

省資源・省エネルギーの実を挙げて地球のシステムや生態系の偏芯を回復するためには、多様で異なる領域的な知による多様な省資源・省エネルギーの営みの間に共通する広域的な知を発見しなければならない。人間の多様な生産・消費・廃棄の活動を互いに矛盾のない連続性のあるものとして実現することが重要である。

2. 省資源・省エネルギーの営みのモデル

省資源・省エネルギーの実を挙げるためには、全ての生産・消費・廃棄の営みが、次の10項目の全てに該当する省資源・省エネルギーの営みとして遂行されなければならない。そのような生産・消費・廃棄の営みを実現するためには、10項目の省資源・省エネルギーの営みのうちのいずれの2つをとっても矛盾なく両立するような広域的な知を発見していかなければならない。

- ①常に新レベルの有意で高性能な財・サービス・知的システムを新たに開発して資源やエネルギーおよび情報の活用効果と利用効率を限りなく高める。
- ②部材の機能の傾斜複合化、財・サービス・知的システムの有意な複合化を図って資源やエネルギーおよび情報の活用効果と利用効率の向上を追求する。
- ③内外の要因に基づくシステムの変動に対し新たな自己組織化によって自律対処のできる誤作動や故障のない安定品質・長寿命の財・サービス・知的システムを生産し消費する。
- ④新レベルの高性能パーツとの交換で当初の性能を上回る若返り修理が可能な財・サービ

ス・知的システムを生産し消費する。

- ⑤ 廃物を燃料としてではなく再利用の可能なパーツや素材として、また、残留熱量を再利用の可能なエネルギー源として、完全に回収し消尽的な循環利用ができるように財・サービス・知的システムを生産し消費する。
- ⑥ 特定の種類の資源やエネルギーおよび情報の大量使用を抑制し解消するため常に代替の資源やエネルギーおよび情報を開発し実用化する。
- ⑦ 高物質・高エネルギー社会の生産・消費・廃棄における摩擦、抵抗、接合や表面の劣化、資源・エネルギー・情報の未利用排出、排気ガス、騒音、振動、粉塵、過剰照明、によるムリ・ムラ・ムダの検出と排除の徹底を期して改善と是正を図る。
- ⑧ 環境に修復困難な影響を与えないように財・サービス・知的システムを生産・消費・廃棄し、損耗を生じた場合は直ちに修復をし保全を図る。
- ⑨ 大気・土壌・海水・河川水・湖沼水・地下水の短期・小域かつ長期・大域の成分分析を通じて人の生産・消費・廃棄の営みと地球のシステムや生態系の偏芯との因果をサベイルランスし原状回復のための暫定・恒久策の徹底を期する。
- ⑩ 人体構成成分の短期・小域かつ長期・大域の成分分析を通じて、地球のシステム・生態系の偏芯と人体のシステムの偏芯との因果、および衣・食・住・その他の生産・消費・廃棄の営みと人体のシステムの偏芯との因果をサベイルランスし原状回復のための暫定・恒久策の徹底を期する。

3. 脳の機能と情報処理の特質

(1) 脳の機能 人の脳の機能は『過去を想起し未来を予期して今ここに対処する』ことである。人は生存環境との間に生じた不均衡を解消するために、不均衡な事態を理解し働きかけるのに必要な新たな情報を取り入れる。そして過去を想起し、新たな情報と過去に蓄積した記憶情報との間の意味の類似と差異を探索する。過去の記憶情報との類似と差異による類比と外挿を行なって、引き続く未来に生起する事物・事象や自らの思考・自らの行動を予期する。未来についてのより可能でより確実な予期を選択しながら今ここに対処している。

(2) 脳の情報処理の特質 知・情・意による脳の総合的な情報処理の機能の特質は『多様で異なる領域的な知の間に、部分と全体の同時最適化を実現して、より広域的な知を発見し創造して実践に移し、生存環境に対する理解と働きかけの連続性を確保しようとする』ところにある。

4. 領域的な知と広域的な知

(1) 領域的な知 領域的な知とは、特定の分野の目的行動を行なうための理解と働きかけをするのに要した、または要すると思われる知の要素および要素の関係とそれらの集合のことをいう。

(2) 広域的な知 広域的な知とは、異なる領域的な知の間に広域的に創発する知のパターンのことをいう。知のパターンとは、脳という認知場において、小域で隣接する3つの記憶の部分域、または中域で近接する3つの記憶の部分域、または大域で離隔する3つの記憶の部分域が、肯定×否定×否定あるいは肯定×肯定×肯定の関係に収束したときに創発するところの2つの記憶あるいは3つの記憶が担う事実や価値や目的の記憶の間の共通のないしは類似的な意味のパターンをいう。3つの記憶の部分域の関係を小域的なもの

から中域的なものや大域的なものに組み換えて、より広域的な知を創発する部分と全体の同時最適化のプロセスは、『3軸認知場』のモデルに『部分域と全体域の誘導合致』のモデルを適用することによって計算論で示すことができる。

5. 『3軸認知場』のモデル

(1) 予期と対処が知のモジュールを蓄積 人は現在から未来に向けて新たな記憶の部分域を形成し、事物・事象・自らの思考・自らの行動に関する事実と価値と目的の情報を、脳という[時間][事実・目的空間][評価(感情)空間]の3軸からなる認知場に、4単位の記憶の部分域が起・承・転・結の順で連なる知のモジュールを積み重ねて蓄積する。領域的な知は、知のモジュールの集合である。

(2) 時間的な知と空間的な知の交叉 知のモジュールは、時間的な知と空間的な知が交叉して形成され、時間的な知・空間的な知・時間的な知の連鎖または空間的な知・時間的な知・空間的な知の連鎖によって成立する。知のモジュールは他の知のモジュールと並存したり、より上位の知のモジュールの構成単位となる。時間的な知とは、時間軸で隣接する独立事象の記憶と従属事象の記憶からなる一対の事実または価値または目的の記憶が空間軸の同じ位置に蓄積されたものをいう。空間的な知とは、空間軸で隣接する独立事象の記憶と独立事象の記憶からなる一対の事実または価値または目的の記憶が時間軸の同じ位置に蓄積されたものをいう。

(3) 想起が広域的な知を創発 脳は現在から過去に向けて新たな記憶の部分域から過去の記憶の全ての部分域へエネルギーや物質を循環させ、新たな記憶の部分域と過去の記憶の全ての部分域を融合という関係で結びつけて、記憶の過去の全てのネットワークと多重させながら、記憶の最新のネットワークを構築する。このプロセスで、3つの記憶の部分域の関係が小域的なものから中域的なものや大域的なものに組み換えられて、より脱領域的で、より広域的な知が創発する。

6. 部分と全体の同時最適化

(1) エネルギー・物質の系

自然のシステムがエネルギーや物質量のレベルの高い部分域から低い部分域へエネルギーや物質を循環させるプロセスにおいて、部分域同士は、ある側面でプラスに関係しエネルギーや物質の移動を促進し合うと共に他の側面ではマイナスに関係してエネルギーや物質の移動を抑制し合い、トータルでは両方の関係が平均されてプラスでもマイナスでもない中立的な『融合』の関係で結合して、部分と全体の同時最適化を実現する。

(2) 情報の系

エネルギーや物質の系としての脳の記憶の部分域AとBがプラスに関係するとき、情報の系としての脳では、記憶の部分域AとBが担う事実または価値または目的の情報が肯定し合って共通的・類似的な意味が見出され、マイナスに関係するときは、否定し合って各々の個別的・差別的な意味が見出される。そして、今ここにおける有意な方の関係が選択されて顕在化する。

新たな記憶情報と過去の記憶情報との共通性・類似性や個別性・差異性が探索されるときに新たな記憶の部分域Aの記憶情報と過去の記憶の部分域Bの記憶情報との2者関係において、肯定ないしは否定の有意な方の関係が選択されるという部分の最適化と、記憶の部分域A・記憶の部分域B・Bよりも過去の記憶の部分域Cの3者の記憶情報の関係にお

いて、否定×否定×否定の関係の組み合わせや肯定×肯定×否定の関係の組み合わせの不安定な関係よりも、肯定×否定×否定の関係の組み合わせ、または、肯定×肯定×肯定の関係の組み合わせの安定な関係が選好されるという全体の最適化が同時に実現するように、脳という認知場が組織化されて、脱領域的で広域的な知が創発する。

7. 想起が広域的な知を創発するプロセスのモデル

脳が新たな記憶の部分域と過去の記憶の全ての部分域を『融合』させて、記憶の最新のネットワークを構築すると共に、より脱領域的でより広域的な知を創発するプロセスは、『部分域と全体域の誘導合致』のモデルによって計算論で示すことができる。

8. 『部分域と全体域の誘導合致』のモデル

(1) モデル式

自然のシステムがエネルギーや物質量のレベルの高い部分域 P_2 から低い部分域 P_1 へエネルギーや物質を循環させるプロセスにおいて、 P_2 と P_1 は、ある側面でプラスに関係しエネルギーや物質の移動を促進しあうと共に他の側面ではマイナスに関係してエネルギーや物質の移動を抑制し合い、トータルでは両方の関係が平均されてプラスでもマイナスでもない中立的な『融合』の関係で結合し、エネルギーや物質の系の部分と全体の同時最適化を実現する。

P_2 , P_1 のエネルギーや物質量のレベルを ϱP_2 , ϱP_1 [$\varrho P_2 > \varrho P_1 > 0$] とし、 $\varrho P_2 = 1$ としたときの ϱP_2 と ϱP_1 の比を、

$$\varrho P_2 / \varrho P_1 = 1 / \varrho P_1 \quad 【1】$$

のように表わすと、 P_2 と P_1 の間のプラスの関係は、

$$\varrho P_2 / \varrho P_1 > \varrho P_2 + \varrho P_1 \quad 【2】$$

P_2 と P_1 の間のマイナスの関係は

$$\varrho P_2 / \varrho P_1 < \varrho P_2 + \varrho P_1 \quad 【3】$$

P_2 と P_1 の間のトータルとしての中立的な『融合』の関係は、

$$\varrho P_2 / \varrho P_1 = \varrho P_2 + \varrho P_1 \quad 【4】$$

というモデル式で表わされる。

(2) モデル項

1つの全体域になろうとする作用の力は、 ϱP_2 と ϱP_1 の格差が大きくなる程に強くなり、格差が小さくなる程に弱くなる。

$\varrho P_2 / \varrho P_1$ の項の値は、 ϱP_1 が小さくなれば増加し、 ϱP_1 が大きくなれば減少するので、 $\varrho P_2 / \varrho P_1$ の項を、1つの全体域になろうとする作用の力を表わすモデル項とすることができる。

各々の部分域のままで留まろうとする作用の力は、 ϱP_2 と ϱP_1 の格差が小さくなるほどに強くなり格差が大きくなるほどに弱くなる。

$\varrho P_2 + \varrho P_1$ の項の値は、 ϱP_1 が大きくなれば増加し、 ϱP_1 が小さくなれば減少するので、 $\varrho P_2 + \varrho P_1$ の項を、各々の部分域のままで留まろうとする作用の力を表わすモデル項とすることができる。

(3) モデル式の解

【1】と【2】からプラスの関係のモデル式の解は、 $\varrho P_1 < (\sqrt{5} - 1) / 2$ となり、【1】

と【3】からマイナスの関係のモデル式の解は、 $\ell P_1 > (\sqrt{5} - 1) / 2$ となる。【1】と【4】から『融合』の関係のモデル式の解は、 $\ell P_1 = (\sqrt{5} - 1) / 2$ となる。

(4) 『融合定数』と『循環定数』

自然のシステムにおいて、エネルギーや物質量のレベルが高い部分域から低い部分域へ順次にエネルギーや物質を循環させて、それらの全ての部分域が多重なネットワークとして融合の関係で結びつくためには、

《1》 $\ell P_2 / \ell P_1 = 1 / \{(\sqrt{5} - 1) / 2\}$ を満たすことのほかに、

《2》 P_2 から P_1 移動するエネルギーや物質量のレベルが、

$\sqrt{\{(\sqrt{5} - 1) / 2\} - (\sqrt{5} - 1) / 2}$ に相当すること、

という2つの条件を満たさなければならないことを、『知の組み換えの基本モジュール』の例によって、計算論で示すことができる。

$[(\sqrt{5} - 1) / 2 \approx 0.61803398]$ を『融合定数』FS [FUSIONAL CONSTANT] と名づけ、
 $[\sqrt{\{(\sqrt{5} - 1) / 2\} - (\sqrt{5} - 1) / 2} \approx 0.168117389]$ を『循環定数』CR [CIRCULATIVE CONSTANT] と名づけることにすると、FSとCRの間には、

$$(FS + CR)^2 = FS \quad \text{という関係が存在する。}$$

9. 『知の組み換えの基本モジュール』

基本モジュールは図のように現在から未来に向けて形成される6個の記憶の部分域F, G, H, I, J, Kからなる。

FG, HI, JKの結びつきは時間的な知であり、GH, IJの結びつきは空間的な知である。

それぞれの部分域の間が肯定の関係の場合は実線、否定の関係の場合は点線で表わされる。

新たな記憶の部分域Kから過去の記憶の部分域J, I, H, G, Fへエネルギーや物質を循環させて融合の関係で結合させ、3組の記憶の関係をKJIという小域的なものからKIHという中域的なものやKHGという大域的なものに組み換えて、KHGというより広域的な知のパターンを創発するプロセスの基本モジュールを10ステップで記述することができる。

[1] エネルギー・物質量の大きさがxのKとエネルギー・物質量の大きさがFS×xのJが融合、KからCR×xが減ってKの値はx(1-CR)になり、JへはCR×xが増えてJの値はx(FS+CR)になる。

この表現を略して、K[x]とJ[FS×x]が融合、Kは[x(1-CR)]に、Jは[x(FS+CR)]になる、と記すこととし、[2]からは略記による。KJは肯定の関係であると仮定。

[2] J[x(FS+CR)]とI[FS×x(FS+CR)]が融合、Jは[x(FS+CR)(1-CR)]に、Iは[x(FS+CR)²]になる。JIは否定の関係であると仮定。

[3] I[x(FS+CR)²]とH[FS×x(FS+CR)²]が融合、Iは[x(FS+CR)²(1-CR)]に、Hは[x(FS+CR)³]になる。IHは否定の関係であると仮定。

[4] K[x(1-CR)]とI[x(FS+CR)²(1-CR)]が融合、Kは[x(1-CR)²]に、Iは[x(1-CR)(FS+CR)]になる。KIは否定の関係となる。

[5] I[x(1-CR)(FS+CR)]とH[FS×x(1-CR)(FS+CR)]が融合、Iは[x(1-CR)²(FS+CR)]

に、Hは $[x(1-CR)(FS+CR)^2]$ になる。I Hは否定の関係であると仮定。

なお、[3]におけるHの $[x(FS+CR)^3]$ は、Hを新たな記憶の部分域とする過去のネットワークに循環し、それを活性化する。

[6] $H[x(1-CR)(FS+CR)^2]$ と $G[FS \times x(1-CR)(FS+CR)^2]$ が融合、Hは $[x(1-CR)^2(FS+CR)^2]$ に、Gは $[x(1-CR)(FS+CR)^3]$ になる。HGは肯定の関係であると仮定。

[7] $K[x(1-CR)^2]$ と $H[x(1-CR)^2(FS+CR)^2]$ が融合、Kは $[x(1-CR)^3]$ に、Hは $[x(1-CR)^2(FS+CR)]$ になる。KHは肯定の関係となる。

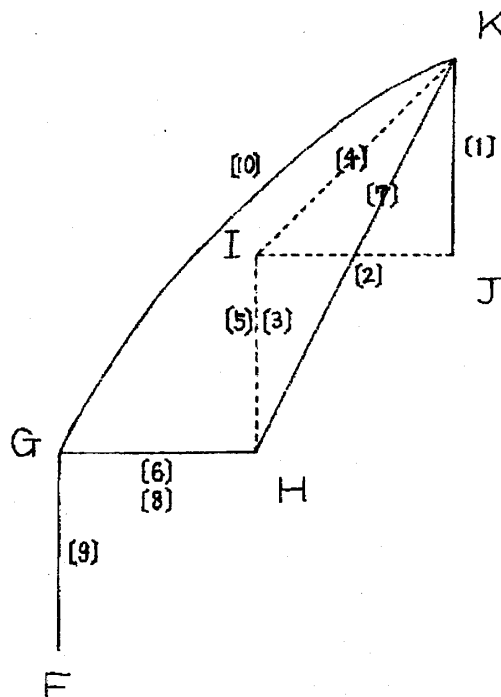
[8] $H[x(1-CR)^2(FS+CR)]$ と $G[FS \times x(1-CR)^2(FS+CR)]$ が融合、Hは $[x(1-CR)^3(FS+CR)]$ に、Gは $[x(1-CR)^2(FS+CR)^2]$ になる。HGは肯定の関係であると仮定。

なお、[6]におけるGの $[x(1-CR)(FS+CR)^3]$ は、Gを新たな記憶の部分域とする過去のネットワークに循環し、それを活性化する。

[9] $G[x(1-CR)^2(FS+CR)^2]$ と $F[FS \times x(1-CR)^2(FS+CR)^2]$ が融合、Gは $[x(1-CR)^3(FS+CR)^2]$ に、Fは $[x(1-CR)^2(FS+CR)^3]$ になる。GFは肯定の関係であると仮定。

[10] $K[x(1-CR)^3]$ と $G[x(1-CR)^3(FS+CR)^2]$ が融合、Kは $[x(1-CR)^4]$ に、Gは $[x(1-CR)^3(FS+CR)]$ になる。KGは肯定の関係となり、KHGという肯定×肯定×肯定の知のパターンが創発する。

以上のように、[4]において3組の記憶の小域的な関係K J Iが成立し、[7]では3組の記憶の中域的な関係K I Hに組み換えられ、更に [10] で3組の記憶の大域的な関係K H Gに組み換えられて、より広域的な知のパターンがコヒーレントに創発するプロセスを通じ、新たな記憶の部分域Kが近接または離隔する全ての記憶の部分域と4次のパスで結合することを確認できる。



図

インターネットを利用した遠隔健康管理システムの試み

○里村 宏章¹, 田中 猛彦², 吉廣 卓哉², 中川 優²

¹和歌山大学大学院システム工学研究科, ²和歌山大学システム工学部

Healthcare Supporting Systems over the Internet

○ Hiroaki Satomura, Takehiko Tanaka, Takuya Yoshihiro, Masaru Nakagawa

Faculty of Systems Engineering, Wakayama University

We have constructed the healthcare supporting systems for aged people, for patients with diabetes, and for pediatric patients and their families. In these systems, the people in need of nursing care actively and regularly send the healthcare data while the medical personnel view them to make use of subsequent medical services or nursing cares. The architecture is almost common to these systems in the sense that there exists a web and database server together with Internet-enable devices. However, since the data in the databases and to be exchanged were entirely difficult, the systems have been designed, implemented and maintained separately. In this paper, we review the features of the systems, and discuss the united, versatile database system for remote healthcare support.

1. はじめに

在宅看護・介護といえは高齢者を対象とした訪問看護をまず思い浮かべるが、その対象は小児や、一見すると健常者にしか見えない人々にも広がりつつある。その背景として、近年の医療技術の進歩により、これまで救命できなかった疾患を持つ者も救命されるようになったため、重度の障害や難病を抱えたまま療養生活を行う若年障害者が増加している点が挙げられる。そのため要介護者の症状は多様化し、介護者だけでなく医療従事者にかかる負担も増加している。

この問題を解決するため、日々の健康情報を要介護者自身が入力するというアプローチでシステムを構築し、高齢者、糖尿病患者、小児患者の健康情報管理に適用している。この開発の過程で、これまでのシステムのそれぞれの機能や特徴を再考することができ、より多様なニーズに対応できる総合的なシステムの可能性を考えることができた。本稿では、そのデータ取りまとめやシステム開発における留意点を報告する。

2. 遠隔健康管理システム

2.1 遠隔健康管理システムとは

遠隔健康管理システムは、コンピュータや携帯電話などの通信機器を利用して要介護者（以下「患者」という）の健康情報を医療機関等に送信し、その情報をデータベースが管理することで、体調管理や病状の発見・予防に役立てるシステムである。インターネットや携帯電話の普及により、様々な遠隔健康管理システムの開発・研究が行われており、潜在利用者数は1,000万人を超えられている[1]。

一般的な遠隔健康管理システムの構成を図1に示す。遠隔健康管理システムの利用者は、情報を主体的、もしくは定期的に発信する患者側と、それらの情報を閲覧し、医療や介護に役立てる

医師や看護師（以下「医療従事者」という）から構成され、両者はシステムを介して情報交換を行う。緊急時には電話などでコミュニケーションをとることもある。

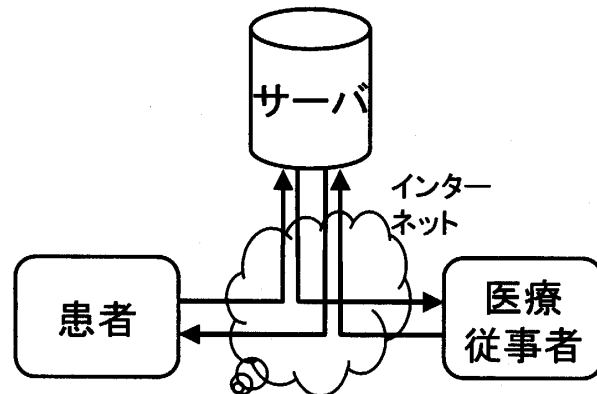


図1：遠隔健康管理システムの構成図

2.2 これまでに開発してきた遠隔健康管理システム

ここでは、筆者らの研究室でこれまで開発してきた3つの遠隔健康管理システムを紹介する。

①花園村遠隔介護システム[2][3]

このシステムは、和歌山県立医科大学看護短期大学部および和歌山県看護協会との共同研究として構築したもので、訪問看護師が要介護者全戸へ十分に訪問するのが容易でない和歌山県花園村（現かつらぎ町）に導入し、2004年2月から11月にかけて運用実験を行った。

回答用機器として、電話機の延長として操作可能なLモード端末を使用した。認証には、Lモード端末固有の番号（固有識別番号）を用いた。質問は、薬剤管理・栄養管理・体調管理に関する8問とし、通話料を考慮して3分以内に回答できるようにした。また本システムでは、どの回答者も共通した質問に答えてもらうものとした。

②糖尿病看護支援システム[4]

このシステムは、1～2か月に一度通院して、アドバイスや薬品の処方を受ける、比較的軽度な糖尿病患者を対象として、日常の健康管理を支援するものである。2005年12月より、和歌山労災病院にて試行実験を実施している。

患者は、貸し出された携帯電話（iモード端末）を用いてサーバにアクセスする。看護師は、病院内のPCからサーバにアクセスして、患者からの報告データを閲覧するだけでなく、患者ごとの登録や療法の設定・変更も行える。

回答内容は、運動療法、食事療法、薬物療法、血糖値報告に大別される。食事療法では、携帯電話のカメラを使って食事画像を撮影し、送信できるようにした。栄養士は、その画像から何を食べたかを把握し、データベースに登録するとともに、患者にアドバイスを送ることもできる。また薬物療法において、使用する薬の服用量やタイミングを誤ることは重大な問題となる。そこで、患者の報告内容に異常があるときに、その内容を色で表示するアラート機能を取り入れている。原因に応じて6種類に色分けして、医療従事者側閲覧画面に表示される。

③小児健康情報管理システム[5]

このシステムは、1か月に一度通院する在宅の小児患者およびその家族を支援するものである。和歌山県立医科大学保健看護学部および和歌山県看護協会との共同研究として構築し、2006年4月より運用実験を実施している。

これまでの2つのシステムと異なる点は、回答者が患者と一致しないところにある。そこで、健康状態を回答するだけでなく、患者の体調や、医師に相談したいことも自由に記述できるようにし、医師や訪問看護師らが返信できるようにした。回答に使う機器は、回答者の使い慣れたPCとしている。

3. 遠隔健康管理システムにおけるデータベースの検討

ここでは、遠隔健康管理用のデータベースを用意することで、様々な要望や仕様に応じたシステムを開発・運用し、多様化する要介護者へ対応できないかという背景で、検討したデータ構造やデータベース設計を紹介する[5]。

3.1 データ構造

今回、患者ごとに質問する項目が異なることを想定している。このとき、患者ごとにテーブルを設ける設計では、患者が増加すると、データベースが非常に煩雑になってしまう。そのため本システムでは、患者専用の単語カードが割り当てられているようなデータ構造をとっている。ここで、説明のために単語カードに例えたデータ構造のイメージを図2に示す。

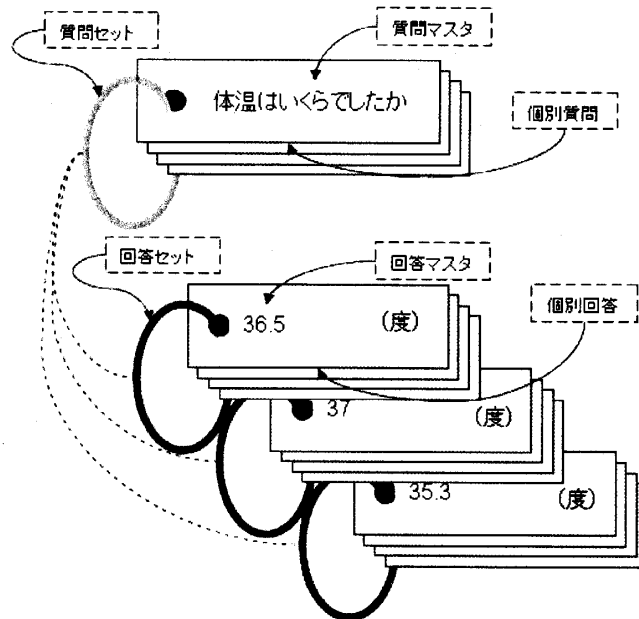


図2：データ構造のイメージ

質問セットは、患者がどのようなタイミングで質問に答えるかという質問全体の情報を統括し、単語カードのリングの役割を果たしている。それに関連して個別質問で質問セットに対応する質問内容を指定する。これは単語帳のカードに相当する。そして質問マスタで質問内容の詳細を格納する。

回答を登録する際には、質問セットに対応した回答セットが回答全体の情報を統括する。そして回答マスタに実際に入力された情報を格納する。また、個別回答には、体温などの数値情報や、体調の良し悪しといった回答内容の詳細が格納されている。

これにより、患者一人に対して個別の質問群を割り当てることができる。そしてその質問群は、様々なシステムから送信されてくる情報を管理することが可能となる。

3.2 データベースの設計

前節で述べたデータ構造に基づき、データベースを設計した。その特徴は以下の通りである。

質問セットと回答セットを1対多の関係として参照キーを設けた。質問セットと個別質問、回答セットと個別回答についても同様である。しかし、個別質問と個別回答、質問マスタと回答マスタには、直接的な参照キーを設けていない。これは、いくつかのテーブルを組み合わせたSQL文を記述することで、どの質問に何と回答した(できる)かが容易に求められるためである。

回答は、選択肢を選ぶだけでなく、体温などの値を記載することがある。データベースにおいては、記述式回答はすべて文字列として扱う。

回答セットにはいつ回答したかに関する、時刻型の属性を設けている。一方、質問セットには、1日のどのタイミング(朝、昼、夜など)で回答すべきかを、時間単位のインターバルとして格納する。

回答内容ごとに色分けして、医療従事者の閲覧画面で識別しやすいようにした。この色情報は、独立した一つのテーブルとして設け、回答マスタが参照する。

4. おわりに

本稿ではより柔軟性に富んだ遠隔健康管理システムの可能性について述べた。現在、筆者らの研究室で3種類の遠隔健康管理システムを開発し、以降高まる需要に対応して更に様々なシステムの開発・運用が予想されるが、根底となるデータベースを一つ設計しておけば、様々な病状を持つ患者や、要望に応じたシステムを円滑に開発、運用することが可能であると考えられる。しかし本案の有効性については未知の部分が多く、本稿で紹介した遠隔健康管理システム以外のシステムへの対応が高いレベルで実現できるかどうかはこれからも検討が必要である。

現在本案は、実証実験を継続中であり、並行して患者、医療従事者から新たな要望や不満点を募って、改良を続けている。

謝辞 システム開発にあたり貴重なアドバイスをいただいた、和歌山県立医科大学看護学部の諸先生方、和歌山県看護協会の皆様、和歌山労災病院の医師・看護師の皆様、および実験に協力いただいた保健師・訪問看護師の皆様および患者様に感謝いたします。

参考文献

- [1] 総務省: 「情報通信白書 平成15年版」, <http://www.johotsusintokei.soumu.go.jp/whitepaper/ja/h15/index.html> (2006年4月14日参照)
- [2] 宮本浩伸; 岡本結宇; 小脇亮平; 吉廣卓哉; 堀内恵美子; 中川優: 「花園村遠隔介護システムの構築」, 情報知識学会誌, Vol. 14, No. 2, pp. 57-60, 2004.
- [3] Yoshihiro, Takuya; Miyamoto, Hironobu; Okamoto, Yuu; Kowaki, Ryohei; Horiuchi, Yumiko; Nakagawa, Masaru: "An Aged Homecare Supporting System in Rural Areas", Proc. Sixth Joint Conference on Knowledge-Based Software Engineering, pp. 293-300, 2004.
- [4] 池本和広; 河村伊津美; 森濱大輔; 吉廣卓哉; 森久美子; 香川幸子; 山本康久; 中川優: 「携帯電話を用いた糖尿病看護支援システム」, 情報処理学会研究報告, 2006-GN-58 (31), pp. 197-202, 2006.
- [5] 里村宏章: 「在宅介護支援のための遠隔健康管理システム」, 和歌山大学システム工学部卒業論文, 22p. 2006.

テキスト入力によるキーストロークダイナミクス

○ 佐村 敏治¹, 西村 治彦²

Keystroke Dynamics in Text Typing

○ Toshiharu SAMURA and Haruhiko NISHIMURA

Biometrics is classified into verification and identification. A lot of researchers of the keystroke dynamics have treated the verification which is used for the user login, based on the assumption that people each type in uniquely characteristic manner. However, its error rate is large compared with other verifications, for instance, by the fingerprint and the retina. In this research, we pay attention to the identification and investigate several characteristics of English text typing on the keystroke dynamics. As an efficient measure, time-interval between press and release of the key (p-r time) is extracted, by which a personal fluctuation of keystroke data is minimized. Experimental results show that the proposed methods on the personal fluctuation are considerably promising.

1 はじめに

人は打鍵(キーストローク)のデータに対して固有のリズムを持っており、この性質を利用することはこれからの情報社会において大きな役割を担っていくと考えられる。キーストロークダイナミクスは打鍵データの特徴を利用したバイオメトリクスである。

キーストロークダイナミクスの研究の多くが利用者のログイン時に利用される照合(verification)認証である[1, 2, 3]。パスワードの認証を知識(記号列)だけでなく、同時にキーストロークダイナミクスも用いて照合認証に利用する。しかし利用者の疲労やキーボードの習熟度に左右されやすく、指紋や網膜などの認証に比べると誤認率または認証失敗率が高くなる。

キーストロークダイナミクスを照合認証のような高い精度が要求される生体認証に用いるのに困難が伴うとしても、筆跡や声紋のような人間の行動に関わるような特徴分析に用いることは価値があるように考えられる。そこで本研究では識別(identification)認証という観点からキーストロークダイナミクスを扱う。パスワード認証のように何度も同じ語を入力するのではなく、全く異なった文章を入力しても個人の特徴が出るような分析方法を提案する。このような自由な文章を入力したときのキーストロークダイナミクスの研究は国内ではほとんど行われていない[4]。

本研究により、クラッキング行為のあった計算機の打鍵データログを解析することでプロファイリングを行ったり、日頃の打鍵データを監視することで個人の健康状態をチェックする医療システムなど多くの応用が期待できる。

最初の取り組みとして平素な英文を入力したときのキーストロークダイナミクスを考える。我々は、論文[5]で、打鍵データにおいてばらつきの生じにくい特徴量として、キーを押している時間(p-r時間)に着目し、その分析方法を議論した。本論文では更に解析方法を改良し、実験についても詳細な解析結果を示す。

2 英文入力における打鍵計測実験

本研究で用いた英文入力の打鍵計測実験について説明する。図1に開発したソフトを示す。本ソフトの開発には、Java言語のJ2SE5.0を利用し、キーの種類と打鍵時間の記録に

¹明石工業高等専門学校電気情報工学科

Electrical and Computer Engineering, Akashi National College of Technology

²兵庫県立大学大学院応用情報科学研究科

Graduate School of Applied Informatics, University of Hyogo

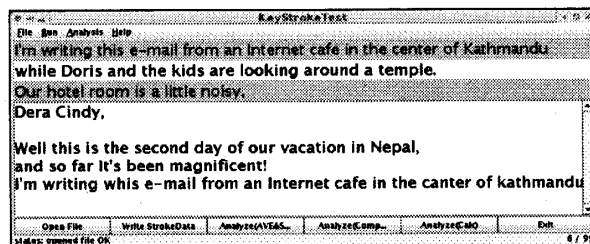


図 1: 打鍵計測実験ソフト (画面例)

は KeyListener を用いた。またキーを押した時刻 (プレス時刻: press time) だけでなく、離れた時刻 (リリース時刻: release time) も記録する。時間分解能はミリ秒である。

英文については英検準2級程度のテキストを用意した。本研究では5名の被験者 (A,B,C,D,E) について実験を行い特徴分析を行った。使用したコンピュータは、CPU Pentium4 1.7GHz, 1024MB の RAM であり、OS は Vine Linux3.2 である。

3 キーを押している時間 (p-r 時間) による特徴量

3.1 p-r 時間の分析

本研究では、大量の打鍵データを扱う中で個人的ばらつきの生じにくい特徴量として、キーを押している時間 (p-r 時間) に注目する。本節ではその理由について説明する。

一口に打鍵と言っても、キーのプレス時刻とリリース時刻を考慮したとき、次のような4つの時間間隔が現れる。

1. プレス時刻とリリース時刻との間隔: p-r 時間
2. リリース時刻とプレス時刻との間隔: r-p 時間
3. プレス時刻とプレス時刻との間隔: p-p 時間
4. リリース時刻とリリース時刻との間隔: r-r 時間

1 は、キーを押している時間 (p-r 時間) と、ある手の指で押してもう一方の手の指で離す時間 (p-r-2 時間とする) の2種類がある。2 は、キーを離して次のキーへ移る時間と、ある手の指でキーを離してもう一方の手の指でキーを押す時間の両方が記録される (r-p 時間)。3 の時間 (p-p 時間) や 4 の時間 (r-r 時間) は、例えば左手であるキーを押した (離れた) あと、すぐに右手で別のキーを押す (離す) ときに記録される時間である。

表 1: 各打鍵時間の平均と標準偏差 (括弧内の数値はイベント数)

	A[msec]	B[msec]
p-r	133.5 ± 27.6 (320)	144.2 ± 31.3 (336)
p-r-2	58.0 ± 37.7 (207)	49.4 ± 32.3 (232)
r-p	138.9 ± 192.3 (372)	127.8 ± 135.8 (410)
p-p	109.7 ± 28.0 (163)	116.3 ± 32.5 (190)
r-r	95.4 ± 33.3 (139)	100.7 ± 44.1 (138)

実際に、被験者 A と B が共通の英文を入力したとき、アルファベットに注目して、それぞれの打鍵時間の平均と標準偏差を表 1 に示す。括弧内の数値はイベント数を表す。

キーを押している時間 (p-r 時間) は標準偏差を見ると個人的ばらつきが少ないことがわかる。一方、r-p 時間はばらつきが大きく、少しの要因で変化しやすい。その他の打鍵時間は標準偏差は小さいが、右手と左手の連携で起こる事象のため、キーボードの経験の少ない被験者は再現性が低い。

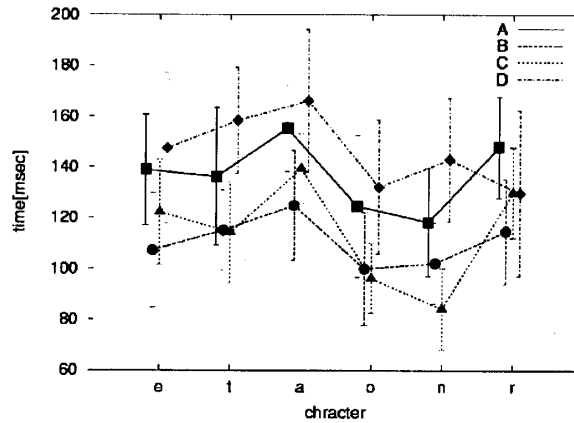


図 2: 各文字 (e,t,a,o,n,r) での p-r 時間の平均と標準偏差

また p-r 時間にどの程度の個人差があるかを調べた。共通の英文を使用した場合、使用頻度の多い文字を対象に各被験者の平均と標準偏差を図 2 に示す。被験者同士で重なる場合も出てくるが個人により特徴が出ていることがわかる。

3.2 p-r 時間の平均と標準偏差を用いた解析方法

本節では p-r 時間を用いた解析方法について説明する。
まず対象とする英文のプロファイルを作成する。

1. 対象とする入力データのうち、“a”～“z”の p-r 時間のデータのみを収集する。例えば、被験者を A とし、アルファベット α (a～z までの文字) の p-r 時間列を

$$K_{\alpha}^A = (k_{\alpha 1}^A, k_{\alpha 2}^A, \dots, k_{\alpha i}^A, \dots, k_{\alpha i_{max}}^A) \quad (\text{ただし}, i_{max} \geq 5) \quad (1)$$

とする。ここで、 i は p-r 時間列の順番を表し、その最大の数を i_{max} とする。ただし、出現回数の少ないデータはプロファイルとして不適当と考えて、5 回以上 ($i_{max} \geq 5$) のデータのみプロファイルとして残す。

2. 式 (1) の平均値 r_{α}^A と標準偏差 s_{α}^A を求める。

$$r_{\alpha}^A = \frac{1}{i_{max}} \sum_{i=1}^{i_{max}} k_{\alpha i}^A, \quad s_{\alpha}^A = \sqrt{\frac{1}{i_{max}} \sum_{i=1}^{i_{max}} (k_{\alpha i}^A - r_{\alpha}^A)^2} \quad (2)$$

次にテストデータとして、被験者 B のアルファベット β (a～z までの文字) の p-r 時間列を

$$K_{\alpha}^B = (k_{\alpha 1}^B, k_{\alpha 2}^B, \dots, k_{\alpha j}^B, \dots, k_{\alpha j_{max}}^B) \quad (3)$$

とする。このとき被験者 A のプロファイルとテストデータとを次式で比較する。

$$N^{BA} = \sum_{\alpha=a}^z \left(\sum_{j=1}^{j_{max}} \frac{k_{\alpha j}^B - r_{\alpha}^A}{s_{\alpha}^A} \right) \quad (4)$$

ただし、プロファイルとの比較対象のアルファベット α は式 (1) で $i_{max} \geq 5$ の条件を満足するアルファベットのみである。以下、式 (4) で得られる値を N 値とよぶ。

3.3 解析結果

解析結果について述べる. 被験者 (A,B,C,D,E) において, まず 3 つの英文 (ここでは, doc1(180 語), doc2(204 語), doc3(188 語)) を入力した. 括弧内の数値はワード数を示す. また, 被験者 A, B, D については別の日に (doc4(109 語), doc5(303 語)) を入力し, 被験者 D には, 更に (doc6(53 語), doc7(55 語), doc8(55 語)) と (doc9(315 語), doc10(316 語)) をそれぞれ別の日に入力した.

被験者 A, 英文 (doc1) をプロファイルとしたときの結果を表 2 左に, 被験者 D, 英文 (doc2) をプロファイルにしたときの結果を表 2 右に示す.

表 2: 被験者 A と英文 (doc1)(表左) または被験者 D と英文 (doc2)(表右) をプロファイルとしたときの N 値

被験者	doc1	doc2	doc3	被験者	doc1	doc2	doc3
A	0.0	-15.4	-33.0	A	-150.9	-188.8	-195.3
B	-860.0	-803.5	-660.4	B	-889.4	-855.4	-719.0
C	-454.9	-449.7	-511.7	C	-525.3	-527.0	-568.7
D	119.7	204.5	237.8	D	-58.0	0.0	4.7
E	-414.9	-191.9	-392.0	E	-556.7	-415.0	-527.3

N 値は 0 に近いほどプロファイル (本人) に近い. 実際の実験 (表 2) によっても本人は他人と比べて N 値が小さくなるのがわかる.

また式 (4) から, N 値が負の場合は, テストデータの方がプロファイルデータより平均 p-r 時間が短く, N 値が正の場合は, プロファイルデータの方がテストデータより平均 p-r 時間が短くなるのがわかる. 従って表 2 の右表から被験者 D の N 値は, 他の被験者との比較で全て負になることから, キーを押しているのは被験者 D が一番長いことが結論づけられる.

以上の分析は, 単位時間あたりの語数でも議論することができる. しかし人の行動の中で, 疲れたり考えごとをしたりして指が止まったりしてしまうと, 単位時間あたりの語数は大きく変化してしまう. 一方で本研究の特徴量を用いると, 実際にキーを押している間しか計測しないので, 指が止まっているような動作は影響を受けない.

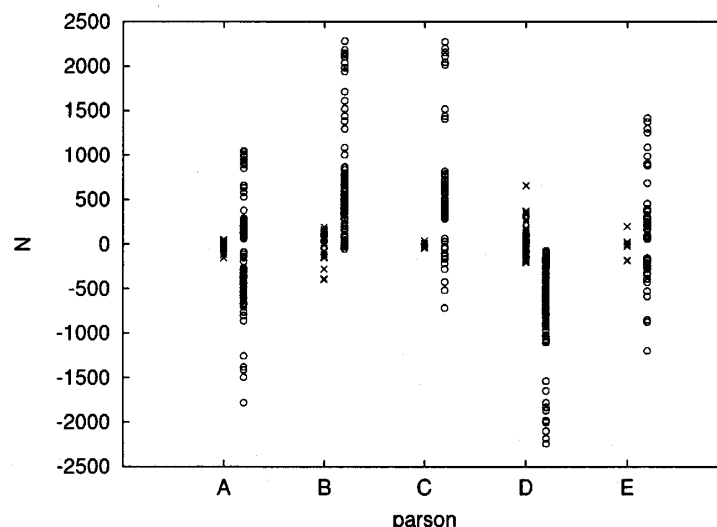


図 3: 本人 (×印) と他人 (○印) による N 値の分布

次に、本人と他人では N 値がどのような分布になるかを図 3 に示す。横軸が被験者、縦軸が N 値を表す。左の×印が本人、すぐ右の○印が他人による N 値である。本人による N 値の多くは $|\Delta N| \sim 300$ の範囲内に納まることがわかる。一方で他人の N 値は広範囲にわたっている。また被験者 B の場合は他人の N 値が全て正なのでキーを押している平均時間が一番短く、被験者 D の場合は全て負なのでキーを押している平均時間が一番長いことがわかる。この解析は、数人の被験者から一人を絞りこむときの方法として有効である。 $|\Delta N| = 100$ を認証の閾値としたとき、認証失敗率(本人ではあるがプロファイルとの比較で $|\Delta N| > 100$ となる確率)と誤認率(他人がプロファイルとの比較で $|\Delta N| < 100$ となる確率)は表 3 になる。

表 3: $|\Delta N| = 100$ を閾値としたときの認証失敗率と誤認率

被験者	認証失敗率 (%)	誤認率 (%)
A	8.0	4.7
B	44.4	10.4
C	0.0	10.1
D	33.3	1.9
E	44.4	6.3

結果として被験者により個人差があることがわかる。また認証失敗率と誤認率とはトレードオフの関係をもつ。1 章でも述べたが、指紋や網膜と比べると精度は悪くなり、本手法を表 3 のような照合認証に用いることは難しい。しかし、以下に述べるように、筆跡や声紋のような人間の行動に関わる分析に用いることは有効な手段となる。また、様々な要因による解析を組み合わせることで、精度のよい認証も可能になると考えられる。この点については今後の論文で述べていきたい。

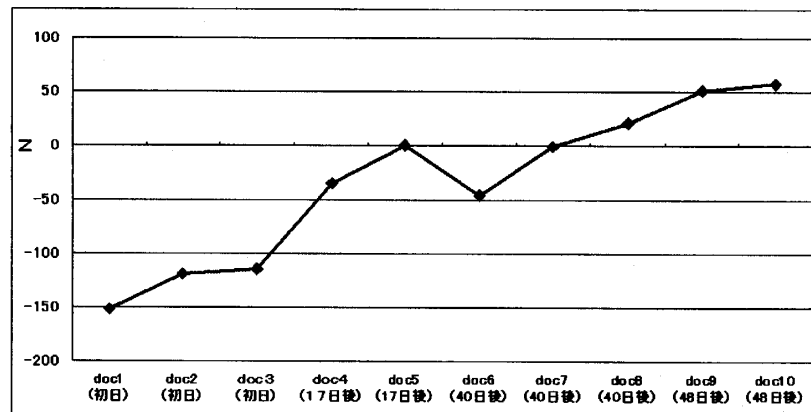


図 4: 被験者 D による N 値の時間依存性 (英文 (doc5) をプロファイルとする)

N 値の時間依存性を調べるために、被験者 D を長期に渡って実験した。図 4 に 英文 (doc5) をプロファイルとして、48 日間に 4 回行った実験結果の N 値を示す。全体的に右上りのグラフになり、 N 値が徐々に上昇している。これは実験の操作及びキーボードの使用に慣れてきて、キーを押している時間が徐々に短くなっていることが理由だと考えられる。この結果を利用すると、長期の観測を行って、ある時に著しい N 値の変化が見られるときは、被験者に心身的な変化(または別人による操作)が影響していると判断できる。

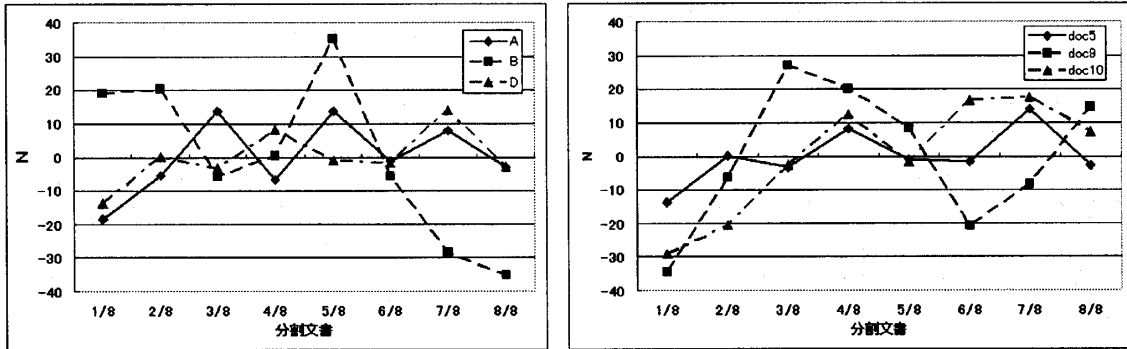


図 5: 英文 (doc5) を 8 分割したときの被験者 A,B,D の N 値の変化 (図左) と被験者 D において英文 (doc5,doc9,doc10) をそれぞれ 8 分割にした N 値の変化 (図右). 分割前の英文をプロファイルとする

最後に、ひとつの文書を分割したときの N 値の変化を求めた。図 5 左は、被験者 A,B,D が入力した英文 (doc5) を 8 分割し、分割前の英文をプロファイルとして N 値の変化を調べた。例えば被験者 B は途中で N 値が増加しているが、徐々に N 値が減少している。長文であるので被験者の疲労または注意力の散乱が関係していると判断できる。

一方、同じ被験者 D に対する比較的長文 (300 語程度) の英文 (doc5,doc9,doc10) を分割して N 値の変化を調べた (図 5 右)。結果として $|\Delta N| < 30$ の狭い範囲で、起伏はあるが比較的安定であることがわかる。以上のような分析により、被験者の心身的な変動を探る手段になることが期待できる。

4 結論と考察

本研究では平素な英文を入力したときに個人らしさを表す特徴量として、キーを押している時間 (p-r 時間) に着目した。そして、この p-r 時間は個人的ばらつきが生じにくいことを示した。次にこの打鍵データの解析方法を提案し、実験による検証を行った。

また N 値の時間的な変化や英文を分割したときの变化も調べた。このような解析は被験者の心身的な変化を分析する手段としてに役立つと考えられる。また、p-r 時間の他の分析方法を組み合わせることで、精度のよい認証も可能になると考えられる。この点については別の機会に議論したい。

本研究のような文章入力によるキーストロークダイナミクスは国内ではほとんど行われていないので今後様々なアプローチから取り組んでいく。

参考文献

- [1] 佐村敏治, 高岡沙緒里, 柴田千恵, 西野順二, 小高知宏, 小倉久和, 打鍵データの特性を生かした個人認証システム, 福井工業大学研究紀要, 29号, pp. 305-312 (1999)
- [2] 粕川正充他: 打鍵データに基づく個人認証システムの評価と改良, 情報処理学会論文誌, Vol. 33, No. 5, pp. 728-735 (1992)
- [3] M. S. Obaidat and B. Sadoun, Keystroke Dynamics Based Authentication, Biometrics: Personal identification in networked society, A. Jain, R. Bolle, S. Pakati (dir.), Kluwer Academic Publishers, pp. 213-225 (1999)
- [4] D. Umphress and G. Williams, Identity verification through keyboard characteristics, Int. J. Man-Machine Studies, Vol. 23, pp.263-273 (1985)
- [5] 佐村敏治, 西村治彦, キーストロークダイナミクスにおける英文入力の特徴分析, 第 50 回システム制御情報学会研究発表会講演論文集, (2006.5)

建築とドキュメントについての考察

山本隆彦（一級建築士）

A Study of relations of the architecture and the document

Takahiko Yamamoto (architect)

The problem of killer condo (fake quake-resistance condominium) struck close to home to an architectural industry and real estate agents.

Various reports and proposals were made public to this problem. Now, a concrete improvement idea is discussed as a prevention plan. However, the content to refer the document of architectural information is few.

In this paper, the current state of an architectural industry and the real estate agents that doesn't try to leave the document positively is described. In addition, it is described that the existence of the document and the establishment of the traceability control this problem.

1. はじめに

建築士による構造計算書の偽装、その偽装によって耐震強度が不足したマンションが多数造られてしまったという、2005年11月17日の国土交通省からの発表は、建築業界や不動産業界に激震として走った。2006年に入り、他の建築士による新たな偽装の疑いが出てきたことで、今回の偽装問題は沈静化するどころか、ますます混迷の度合いが深まった。偽装問題は、建築行政の骨格に位置する建築確認に対する不信感も国民に植え付けた。さらに、構造計算書が偽装された建物に対する現地調査を通じて、設計段階での偽装だけではなく、施工段階での瑕疵が指摘されるに至り、今日の建築生産の仕組みの中に構造的な問題が潜んでいるかのように、建築業界全体が疑惑の目で見られるようになった。

その後、様々な報告や提言が公表され、再発防止策として具体的な改革案や法律の改正案も現在検討途上にある。職業倫理の向上、瑕疵担保責任保険の充実、罰則規定の強化を含む建築行政上の仕組みの再構築などが検討の遡上に挙がっている。

しかし、これら公表されているものを見渡してみても、建物に付随する情報(以下、ドキュメントと呼ぶ)が持つ存在意義について積極的に焦点を当てたものは少ない。ただし、日本建築学会の中間報告では、生産プロセス情報の透明化の実現に向け、建築生産のプロセスにかかわる情報の記録・保存・開示の必要性について言及している[1]。

今回の偽装問題の発生には、ドキュメントを軽視するような風潮も影響している。民間建築の場合、確認申請などの行政の関与や建築士などの権威付けられた資格制度の存在によって、消費者(新築当時の建築主や発注者だけではなく、現在の建物の所有者や居住者)に対し必要なドキュメントが開示されることなく、結果として、消費者から建築生産とい

うブラックボックスの中を覗く手段を奪ってきた。

設計や施工に関わる記録を、生産者(設計者と施工者)と消費者だけではなく、建築確認を担う機関、瑕疵担保責任保険を提供する保険会社、不動産証券化なども視野に入れた場合は金融機関、これらの関係者がドキュメントを長期に渡って保管しかつ必要に応じて内容をトレースするという考え方が、建築業界や不動産業界に根付いていれば、今回の偽装問題は防止できたのではないか。

本稿では、ドキュメントを積極的に残そうとしない建築業界や不動産業界の体質について考察し、さらに、ドキュメントの存在とトレーサビリティの確立が今回のような偽造問題の新たな発生の抑止につながるという考え方についても言及する。

2. ドキュメントに対する姿勢

建築士と建築確認と請負契約、この三つがそれぞれの役割を果たすことを前提として、今日の建築生産の仕組みは成り立っている。この仕組みの始まりは、1949年に建設業法が成立し、1950年に建築基準法と建築上法が成立した時に遡る。

法律での弁護士、会計監査での公認会計士、医療での医師のように、一定の権威を持ち職業倫理にもとづき業務を遂行する資格として建築関係では、建築士が位置する。この建築士の仕事内容を公的に確認する仕組みとして建築確認がある。建築確認とは、建物を建築する際に必要な行政上の手続きのひとつで、その建物の敷地、構造、設備などが建築基準法や都市計画法などの法令に適合しているかどうかを、公的な建築主事(都道府県および政令指定都市に置かれる行政機関)が確認することをいう。なお、建築主事とは別に、国土交通大臣または都道府県知事によって指定された指定確認検査機関と呼ばれる民間機関もあり、ここを使っても、建築確認を受けたことになる。

建築上と建築確認の存在によって、法令が醸し出す権威と専門家としての職業倫理を根拠として、完成した建物の安全性(耐震性)や品質は当然一定の基準を満足しているものとして、消費者は無意識のうちに安心を感じてきた。

また、請負契約とは、建物の完成を目的とした契約である。日本では殆どの場合、施工一括請負方式または設計施工一括請負方式で建物が造られる。すなわち、特定の建物とその建物を施工した建設会社を結び付けることは容易である。

このように、建築上と建築確認と請負契約、この三つの存在を前提した今日の建築生産は、一方では権威と職業倫理もう一方では建築会社の対応という世界を前提としてきた。このような世界では、建物さえ問題なく完成すれば、法令に沿った必要最低限のドキュメントに一部不備があっても特段問題にはならないし、必要最低限を超えるドキュメントを、殆どの場合、誰も要求しなかった。

3. ドキュメントの保管状況

建築確認で法令に適合していると判断されれば、建築確認通知書が発行される。この通知書を受け取って初めて、建築主は着工することができる。構造計算書とは確認申請で提出する書類のひとつであるが、建築確認通知書という「権威」によって、構造計算書という「証拠」(ドキュメントの一部)の存在は無視されてしまう。この構造計算書を含め建築確認の書類とは、設計段階での生産情報である。建築主や発注者が保管しない限り、これ

らの書類は最大5年を超えて保存されることはない。

施工段階での生産情報、例えば現場で用いた施工図、施工記録や検査記録、工事写真なども、工事監理者(建築士)による工事監理や建築主事などによる完了検査が済めばその目的が終わり、これらの生産情報の大半は破棄される。基本的に建築主や発注者に渡されることはない。言い換えれば、完了検査後に交付される検査済証という「権威」によって、施工に関わる様々な「証拠」(ドキュメントの一部)の存在は無視されてしまう。最終的に、建築主や発注者が受け取るドキュメントとは、通常、検査済書を含む竣工図書[2]だけである。

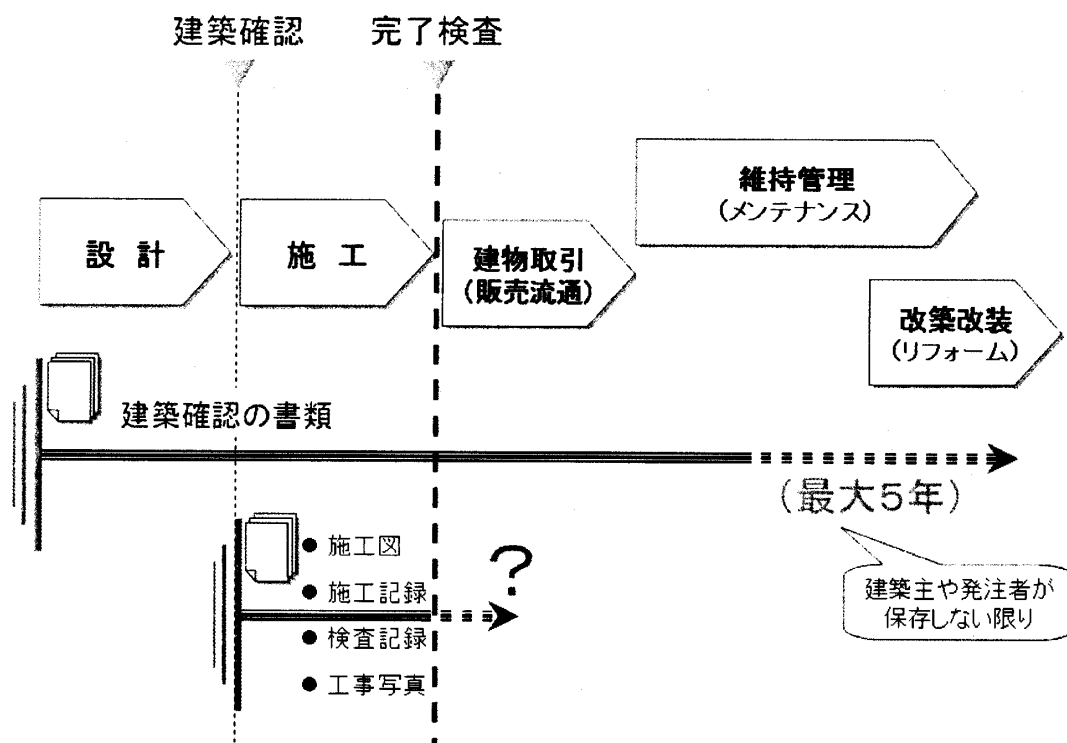


図-1 ドキュメントの保管状況

建物は残るが、建築(設計と施工)の状況を具体的に追える書類はドキュメントとして残らない(誰も要求しない)ことが多い。将来、建物が転売される際に建築確認の書類や竣工図書が紛失していると、これらのドキュメントは新たな所有者へ引き継がれない。

4. 犯罪機会論

『検査を依頼したら、偽装した物件が通ってしまったので、その後も依頼を続けてしまった』と、チェックが甘い指定確認検査機関を使い続けた事実を、構造計算書を偽装した建築士は認めている。これだけを読むと、偽装した建築士とチェックの甘い指定確認検査機関だけの問題であるかのように思ってしまう。しかし、これを犯罪機会論[3]の視点から見ると、まったく違った面が見えてくる。

これまでは、犯罪が起こった原因を犯罪者の人格やそれを生み出した境遇(家庭・学校・会社など)にあると考え、犯罪を防止するための対策も、それを除去(人格の再教育や境遇の改善)することに重きをおいて進められてきた。しかし、結局、これだけでは再犯率を下

げることができなかつた。犯罪の予防に対し、従来とは根本的に異なつた、新しい視点を与えたのが「犯罪機会論」である。大阪小学校内児童殺傷事件の公判で、加害者は『門が閉まっていれば入らなかつた』と述べた。この加害者には心の病があつたと思う。しかし、ふとした気の緩みだけでは犯罪は起こらない。犯罪を誘惑するような環境がそこに存在したから犯罪は起こつたと考え、その環境(例えば、犯罪を誘発し易い場所)を分析し、いかに犯罪者にとって都合の悪い環境を作りだすことが犯罪の防止につながるかを、この犯罪機会論では論じている。

構造計算書を含め、建築の状況を証明する書類など「証拠」(ドキュメントの一部)が残らないということは、犯罪機会論からすれば、「監視性」が低いといえる。監視性が低いところで犯罪は起こる。監視性を高めるには、ハード的な要素として「無死角性」(建物の生産過程を明らかにすること)を高めることと、ソフト的な要素として「当事者意識」(建物の生産過程を自分の問題として監視すること)を高めることの両輪が機能することが必要である。

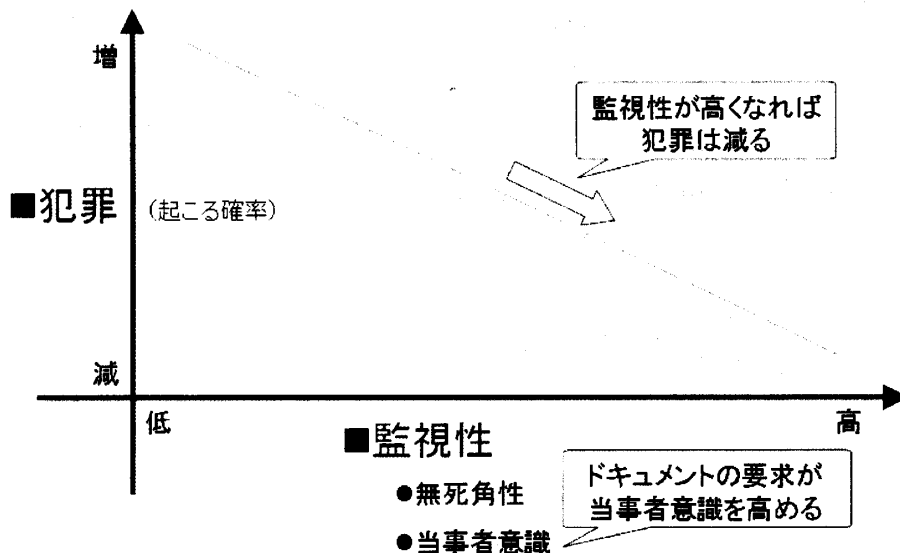


図-2 監視性と犯罪の関係

再発防止策として検討されている罰則規定の強化を含む建築行政上の仕組みの再構築などは「無死角性」を高める。同時に、消費者だけではなく保険会社や金融機関が建築の状況を証明する書類など「証拠」(ドキュメントの一部)の価値を認識しそれを要求するようになれば、「当事者意識」が高くなる。「無死角性」と「当事者意識」の両輪が高まれば、「監視性」の低さは必ず改善される。

5. おわりに

論者は、以前、施工側の視点から、民間建築工事の電子納品(電子データ化された竣工図書の納品)の標準化について検討した[4][5][6]。生産者(設計者と施工者)が自ら手掛けた建物について、ドキュメント(証拠となる詳細な施工記録ではなく、最低限と言える竣工図書)を積極的に残そうしない現状を危惧したからであった。竣工図書の内容の検討だけではな

く、竣工図書をまとめ、その写しを社内で保管するといった一連の作業に対する目的を明確にすることも試みた。そのために、『情けは人の為ならず』との標語を掲げ、電子納品した電子データは建設会社にとって今後の営業の糧になるとのストーリーで意識付けすることを考えた。建築主や発注者に渡すという営業方針を掲げれば、工事現場から完成度の高い竣工図書が届き、社内で自動的に蓄積されると構想したからであった。しかし、消費者側からの積極的な要望も無かった状況下で、生産者側だけでこのようなことを検討すること自体に、当時としては無理があった。

しかし、今回の偽装問題を契機として、今後は、消費者側から、建物に関する様々なドキュメントの開示要求、それもトレーサビリティ(遡及可能性)を確保した上での要求が出てくることだろう。業界の体質改善のためにも、生産者はこの要求に応えなければならない。この要求に対し生産者として適切に対応することを通じて、業界内部においてもドキュメントに対する見方が変化し、長期的に見れば建物の品質も徐々に高まっていくと考える。その時には「竣工図書提出」から「建物情報開示」へと表現も変わっているだろう。

最後に、「建物情報開示」の対象となるドキュメントの案を、以下に示す。

- ① 既存の設計完了時点のドキュメント：確認申請書の副本、設計図書
- ② 既存の建物竣工時点のドキュメント：一般的な竣工図書[2]
- ③ 設計状況を説明できるドキュメント：建築主意図の文書(ブリーフ)、設計説明書
 ※ 建築主意図の文書(ブリーフ)[7]とは、要求する建築の内容を設計者に明示した文書。設計説明書[8]とは、設計図書に表現されない設計と条件、設計意図やコンセプト、設計基準のグレードなどをまとめた文書。
- ④ 施工状況を説明できるドキュメント：施工図、施工記録、検査記録、工事写真
 ※ 中間や完成検査(建築確認業務の一部)や工事監理(建築士が担う業務)で提示した資料の一部

①②のような旧態依然とした確認申請書の副本、設計図書、竣工図書だけではなく、③④のような新たなドキュメントも必ず対象にすべきである。

これらの開示図書(ドキュメント)が情報として建物に長期に渡り付随するようになって初めて、建物情報のトレーサビリティの確立に向けた取り組みは始まるのである。

参考文献と注記

- [1] 日本建築学会 健全な設計・生産システム構築のための特別調査委員会：「健全な設計・生産システム構築のための提言の枠組み(2006年3月9日)」,
<http://www.aij.or.jp/jpn/databox/2006/060310-2.pdf>(2006年4月8日参照)
 「3 提言の要旨」の「A 設計・生産システムを自ら改善していく仕組みづくり」の「A-3 生産の仕組みの改善」の「A-3-3 生産プロセス情報の透明化」参照
- [2] 日本建築学会編：「建築工事標準仕様書・同解説 JASS 1 一般共通事項」, 日本建築学会, 2002.1
 「7節 引渡し」の「7.2 引渡し」より引用
 『完成検査に合格して工事の目的物を発注者に引き渡す場合、監理者と協議のうえ、以下の各種の書類・物品またはこれに代わる目録を添えて、引渡しを行い、必要に応

じて発注者に説明する。その後、発注者が建物・設備を適正に運用・運転できるように協力する。

- (1) 完成届けおよび引渡書
- (2) 鍵・工具引渡し書および鍵・工具
- (3) 特記による予備材料および物品
- (4) 建物取扱い説明書
- (5) 契約で定められた図書 』

[3] 小宮信夫 : 「犯罪は『この場所』で起こる」, 光文社新書, 2005.8

犯罪には「したくなる」環境と「あきらめる」環境があると説く犯罪機会論は、犯罪に対する「原因追及」の呪縛を解き、犯罪の予防を新しい視点から捉えた犯罪学。

小宮信夫(立正大学文学部社会学科助教授)の Web サイト :

<http://www.ris.ac.jp/komiya/>(2006年4月8日参照)

[4] 山本隆彦 他 : 「建築生産における竣工図書の電子化に関する研究-1」, 日本建築学会第17回建築生産シンポジウム論文集, 2001.7

[5] 山本隆彦 : 「『(仮称)民間工事における竣工時に提出する図書(電子データ版)ガイドライン』策定に向けて」, 日本建築学会第1回建築情報化技術フォーラム, 2002.2

[6] 山本隆彦 他 : 「建築生産における竣工図書の電子化に関する研究-2」, 日本建築学会第18回建築生産シンポジウム論文集, 2002.7

[7] 国土交通省 構造計算書偽装問題に関する緊急調査委員会 : 「構造計算書偽装問題に関する緊急調査委員会 報告書(平成18年4月)」,

<http://www.mlit.go.jp/kisha/kisha06/15/150406/03.pdf>(2006年4月8日参照)

「第3章 これからの建築社会のあり方に向けた提言」の「(2) 建築主の役割と建築設計システムの改革」の「① 建築主の役割の明確化」より引用

『今回の事件では、偽装した建築士、見過ごした確認機関の対応策が焦点になっているが、建築主は建築生産システムの頂上に位置しており、法に適合し社会的な資産として評価できる建築を造る責務を負っている。その責任を明確にする手段として、建築主は、建築物の内容、性能や機能、敷地・財務などの制約条件や情報を設計者に文書(ブリーフ)で伝達し、記録に残すことが重要である。』

[8] 日本建築家協会編 : 「設計情報の正しい伝えかた(1) 建築実施設計図書作成基準」, 彰国社, 2000.10

「第5章 竣工図書は建物のライフサイクルを支える重要情報」の「5-1 竣工図書」の「4) 設計説明書」参照

※ 本稿は、「連載：住のトレーサビリティ 建築とドキュメント」(月刊IM 2006年4月号～6月号)、「旅隆人の A/E/C NEWS」(CAD&CG マガジン 2006年2月号～4月号)での発表をもとに、内容を見直し、加筆したものである。

歴史から洞察される判断基準としての観点

○安平哲太郎、佐脇政隆

独立行政法人 産業技術総合研究所 技術情報部門

A view point as a criterion of judgment seen into from our history

Tetsutaro Yasuhira, Masataka Sawaki

National Institute of Advanced Industrial Science and Technology

Office of Technology Information

Abstract : Our object that we see into the future of our society is to measure the distance between the future society and the present condition and create things necessary for leading to the future. Therefore the future society seen into must be inevitable, that is, if we want to exist, we must head for it, able to be co-existent and have a high regard for the self-determination right that the people in the future can construct their society.

Such the future society is described as view points which are effective as a criterion of judgment when selecting measures necessary for co-existence. Such view points appear as the direction in which we must resolve the social problem or which is indicated explicitly by the social insight which means part of people in our society exhibit their insight regardless of notables or not. A view point which works daily in the future as a criterion of judgment is a fusion of view points which have continued since social problems and insights and we call it a view point as a criterion of judgment in the future seen into from our history because our history just consists of social problems and insights arranged on line of time. Moreover we describe how to see into these view points from our history.

Finally we propose the system that a citizen can see into a view point in the future from our history, inspect it during a free and voluntary daily life and recognize it in common in our society.

1. はじめに

未来に関して、予測する手法は幾つか開発されてきた。これは様々な前提条件を用いて現在の延長として未来を考える手法¹⁾である。それに対して、予知やチャンス発見といわれる手法^{1,2)}が考え出されてきた。これは、現象の中にある非線形性(チャンスの芽)を考慮するもので、シナリオを用いて社会の中で潜在化しているチャンスをチャンスとして意識化し、必要な行動をとろうとするものである。これは地震予知や市場の開拓に用いられている。また、庄司³⁾らは、チャンス発見のためにはコンセプトの精緻化が極めて重要である事を指摘した。しかし、これらの方法では、未来を想像する際に指導原理がない為に様々な方向が出てきてしまう。また、未来社会の姿を想像し実現すると考える事によって、その時代の人達はその時代の社会を構築するという意味での自決権との関わりがはっきりしない。また、社会の中に芽生えてくる新しい芽だけから想像される未来像では、過去の

失敗を未来に活かそうという考えが含まれた未来像とはいえない。

我々は、もし生存しようとするなら向かわなければならぬ未来社会がわかれば、我々は現在どこにいるかがわかり、そこへ向かうための準備をすることが出来ると考え、その様な未来社会とはどの様なものであるかを考察した。その結果、社会の問題や社会の一部の人達が洞察力を発揮した現象には、人類共存に向けて社会の問題を解決し得る、あるいは、洞察力が明示的に示す人類共存の為の方向が存在していて、それは我々が日常生活のための手段を選択する際の判断基準となっている事が分かった。そこで、個々の社会の問題や洞察力に付随して現れる方向を今後判断基準としての観点ということにする。さらに有名、無名に関わらず社会の一部の人達が洞察力を発揮した現象を社会の洞察力という事にする。そして、時系列的に発生した社会の問題や社会の洞察力を契機として判断基準としての観点が始まり、それらが継続して未来において融合し、あるテーマ（環境問題、組織と個人、人生等）に関して人類共存と繁栄のために判断基準として意味を持つ時、それを「歴史から洞察される、未来における判断基準としての観点」と呼ぶ事にし、その様な観点で未来社会像を記述する事にする。したがって、チャンス発見はこのような社会の問題や社会の洞察力が現れる予兆を発見するための研究であり、我々は主として歴史上の社会の問題や社会の洞察力が生じた後の向かうべき方向を研究の対象とするわけである。

現在、科学技術の進展に伴い、高度でしかも一般社会に重大な影響を与えかねない（遺伝子技術の進展による疾病因子の除去に伴い人類の多型性が失われるとか、再生医療が脳にまで及ぶとアイデンティティーの問題に直面する等の）懸念が出てきている⁴⁾。これらは、まだ予想の段階で、社会の問題という程の事態を引き起こしているわけではないが、社会の問題がおきてからでは遅すぎる。このような場合、もし、歴史の流れから継続性を持って歴史の指し示す判断基準としての観点が発見されていて、市民の自由で自主的な日常生活によって十分にその正しさが確認されていれば、それを判断基準として再生医療問題や遺伝子技術による疾病因子の除去問題について解決すべき方向を日常の言葉で考え選択してゆける筈である。この事からも、「歴史の指し示す未来における観点」を発見する方法論を実践を通して確立する事が急務であり、その第一歩としての方法と実践法を提案する。

2. 未来社会像の条件

未来社会像を描く時に問題になるのは、それに対する逆読みである。すなわち、それを知った人たちが、現在の自分に不利益であると考えた場合にそういう方向へ向かわないように行動パターンを変えてしまうからである。しかしながら、逆読みが必要な場合もある。たとえば、未来社会像が人類の未来の生存を脅かすような事態の場合、我々はそのような将来像を支えている条件を明らかにし、そうならないように条件を変えてゆくからである。この事は未来社会の具体的な有り様は、基本的にはその時代の人々が共存できるように決定するという事でもある。

はじめにでも述べたように、我々が未来社会像を洞察するのは、現状と我々が向かうべき未来社会との距離を知り、未来社会に至るためにはさらに何が必要であるかを考察し、創造する為である。逆読みが行なわれる事によって変えなければならぬ未来社会像では困る。したがって、上の考察から、未来社会像を描くにあたって必要な3つの条件のある

事がわかる。1 番目は、人類共存のための未来社会像なら逆読みは必要ないと言えるし、2 番目は、未来社会像が我々の思惑や意図を超えて必然（すなわち、我々が生存しようと思うならそこへ向かわざるを得ない社会像）であれば、逆読みに強いと言える。そして、3 番目は、未来の人たちがその時代の社会を構築できるという自決権を尊重したものでなければならぬという事である。

3. 社会の問題、社会の洞察力と判断基準としての観点

ここでは上に述べた3条件を満たす未来社会像を考察する。そのためにまず、産業育成と言う観点から公害問題を通して環境保全という観点が必要となり、さらに環境保全を考慮した産業育成とかわっていった歴史過程について考察する。図1に示すように、明治に渋沢栄一によって産業育成という観点が提起されて以来、産業育成が最重要視されてきた。1970年頃から公害が世間に広く知られるようになり、公害問題が社会の問題となり始めた。その結果、しばらくは従来の観点を判断基準とする人たちと、新しい観点を判断基準とする人たちの間で対立が生ずるが、やがて、新旧観点が融合して環境保全を考慮した産業育成が観点となり対立が解消する。この事は、社会の一部の人たちの洞察力によって、あるいは、社会の問題を解決し得る方向として判断基準としての観点が出現し、将来において、新旧観点が融合して新しい観点（日常的な環境保全を考慮した産業育成）となることを意味している。したがって、この様な融合した観点は、社会の問題を解決し、将来において

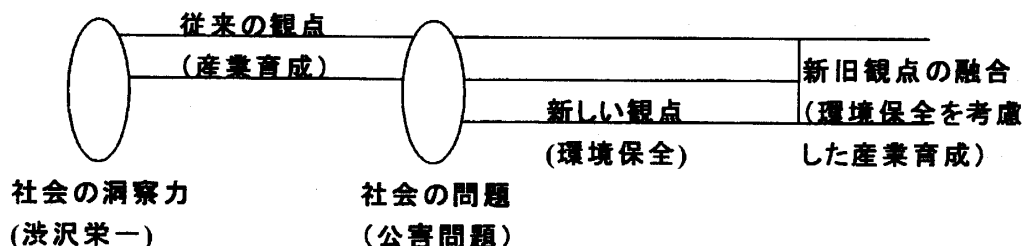


図1 判断基準としての観点

より良く共存する為の観点であるから逆読みは必要ないといえるし、また、現実に我々の手段の選択の判断基準となる観点であるから、その観点に従わざるを得ないと言う意味で逆読みに強いと言える。さらに将来の人たちがその時代の社会を構築する際の判断基準となる観点でもあるから、2. で記述した3条件を満足している未来社会とは、このような融合した判断基準となる観点が記述されたものだという事になる。

4. 歴史が指し示す判断基準としての観点

歴史が指し示す判断基準としての観点を環境問題⁵⁾について考察する。図2において楕円は、社会の問題や社会の洞察力が発揮された状態を表し、その下に記述されているのは社会の問題そのものかあるいは洞察力を発揮した人である。そしてそこから新しい観点が現れ右方向に続いている事を表す。従って、未来においてはこれらの各観点が融合し、日常的に環境保全と社会内循環と社会・自然環境間循環に基づく産業育成が行われる方向、すなわち、歴史の指し示す判断基準としての観点となる。社会・自然環境間循環は、方向性であり実証されているわけではない。けれども、社会内循環がリサイクルさせるために

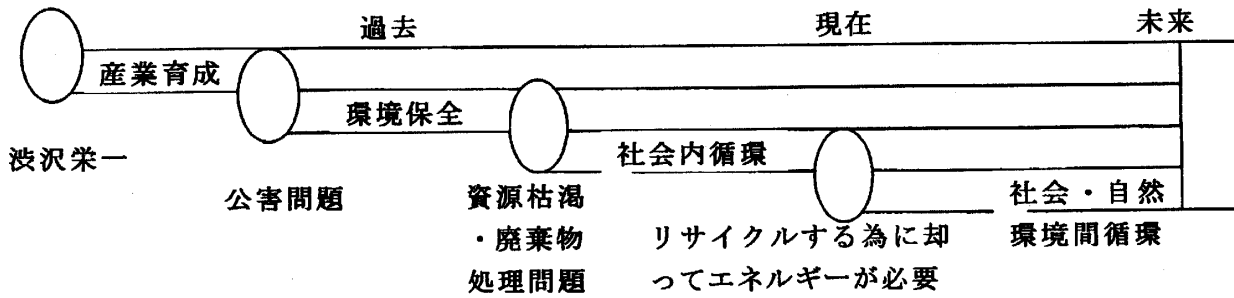


図 2 環境問題における歴史の指し示す観点

エネルギーを必要とするから、自然環境に戻して自然環境のエネルギー（熱、圧力等）を利用すべきであるという事は容易に想像される。この自然環境を通した循環は社会の洞察力に相当するブルントラント委員会が提唱した持続発展可能な開発という見解に対する 1 つの解にもなっている。また具体的事例として、EM菌による生ゴミ処理とその肥料としての利用やバイオマスなどはこれに近い考え方である。

5. 歴史の指し示す観点発見の方法

上記の歴史の流れ図を作成するためには、社会の問題や社会の洞察力が示す現象に対して、社会の問題を解決し、事態をよりよくする為の判断基準としての観点を発見できなければならない。このような発見はコンセプトの精緻化で説明できる。コンセプトの精緻化は日常の購買行動を観測する事によって庄司ら³⁾によって指摘された。基本的には、2つの事象がある観点から見て同じ性質持っていても、両事象が異なるようにさせるためには、別の観点から見て具体的性質が異なるようにさせなければならないという事である。図 3 に歴史の指し示す観点とコンセプトの精緻化との関係を示す。破線部分を見つける事がコンセプトの精緻化である。

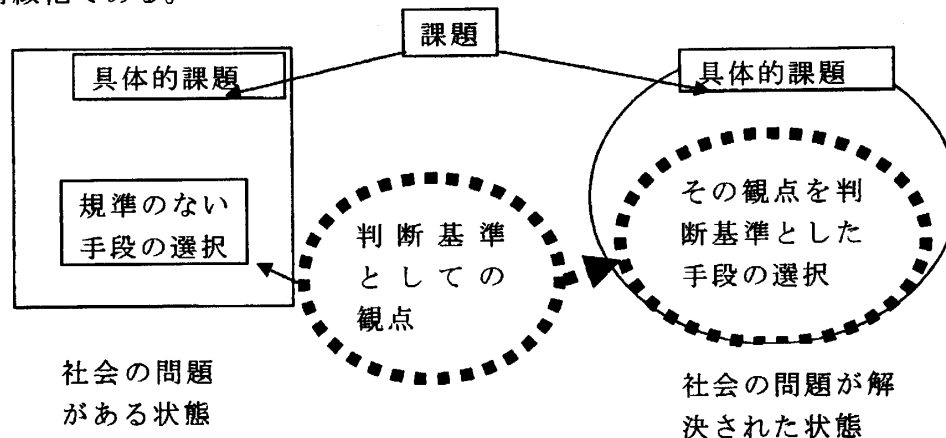


図 3 コンセプトの精緻化と判断基準としての観点との関係

したがって、社会の問題も社会の洞察力の場合も、

- (1) 社会の問題や洞察が生じる前の現象と比較する事によって、問題や洞察の焦点を明確にする事ができる。

(2) それと類似した(身近な、あるいは単純な)現象、あるいは共通部分を持ちながらその現象とは異なる現象と比較する事によって、明確になった問題の本質(その問題が将来における共存と繁栄の為に持つ意味)がわかる。

そして、それを判断基準としての観点とする事ができると考えられる。

5-1. 仮説形成

様々な歴史観がありうる事から、1つの社会の問題に対して様々な観点があると考えられる。従って、現在から過去に渡る具体的な社会の問題や社会の洞察力による発言や提言、あるいは自分の人生の中で経験した特に歴史事象の中に

(1) 人類共存と繁栄のための

(2) 日々の政策や手段の選択の判断基準となりうる

観点があるとしたらそれは何であるかを集団内のブレイン・ストーミングで決定する。それに対して、ブレイン・ストーミングの調停者は、前後の事象や類似した現象と比較して提案された観点を整理し、位置づけてゆく。それらの整理し位置づけられた観点が歴史の中であるテーマに沿って(例えば、経済活動、産業のあり方など)1つの流れが見えてきたところで、人類共存のために指し示す将来の方向を考える。それを「歴史の指し示す未来における判断基準としての観点」とする。さらに、次から次へと生じる社会の問題や社会の洞察力が意味する判断基準としての観点を「歴史の指し示す未来における判断基準としての観点」に位置づけてゆく。

5-2. 仮説の検証

次に我々は、その仮説を検証しなければならない。その仮説は市民の共存のための日常の判断基準に役立つ事を目的としたものであるから、その仮説を参考にした市民の自由で自主的な生活によって、その仮説が日常の判断基準に役立つかどうか検証される事になる。

5-3. 仮説の形成と検証を実現する社会システム

次に仮説形成と仮説検証を実現するシステムについて説明する。この様な仮説は基本的には、市民のブレイン・ストーミングによって立てられるが、前後の歴史事象や、様々な類似したところのある歴史事象と比較し位置づけなければならないし、また、市民による仮説の検証結果を分析し位置づけなければならないという点で専門家の協力は必要である。したがって、歴史専門家の助言を受けた市民による仮説形成集団によって立てられた仮説「歴史の指し示す未来における観点」を社会へ提示し、社会はそれを参考にした自由で自主的な日常生活によってその観点的日常の判断基準としての有効性を検証し、検証結果を提言あるいは問題提起といった形で仮説形成集団に返すという社会と仮説形成集団との間の情報循環⁹⁾を利用する事になる。

重要な事は、

(1) 仮説形成集団は、歴史専門家の助言を得る市民によって組織される

(2) 仮説形成集団で決められた「歴史の指し示す観点」は、あくまでもその集団内の主観に基づく仮説であるという事

(3) その仮説の検証は、その仮説による強制ではなく、その仮説を参考にした自由で自主的な市民の生活による実践によらなければならない

(4) 仮説に対して市民が柔軟に問題提起ができる事である。

すなわち、基本的には市民自身によって「歴史の指し示す観点」に相当する仮説が形成さ

れ、市民自身がその仮説を自由で、自主的な生活によって検証してゆくのである。

6. まとめ

我々の問題意識は、もし、我々が生きようとするなら向かうべき、かつ、人類共存可能な未来社会像を発見出来るならば、その様な未来社会に対して現在どの位置にいるかが分かり、目標に対してさらに何をすべきかがわかるはずという事であった。考察の結果、それは社会の問題を解決する、また、社会の洞察力が明示的に示す方向であり、それは未来の人達がその時代の社会を共存に向けて構築してゆく際の手段の選択の判断基準となる観点でもある事がわかった。

吉岡齊は「こうでなければならぬ」の提案は、必ず「実現可能なもの」に基づくべきであるとした⁷⁾。この場合、歴史の指し示す観点は我々にとって判断基準であるから、「ねばならぬ」に近いものであるが、当然現状では実現可能でないものもあり得る。この事については、

(1) 歴史の流れ(社会の問題や社会の洞察力の時間的配列)から将来身につけざるを得ない観点が継続的に提示される事によってその観点にかなりの確度を持たせる事が出来、

(2) 同時に並行する社会の問題や洞察力が作る別の歴史の流れが指し示す観点をも考慮する事によってその観点到具体性を与えてゆく事が出来その事で創造し得ると考える。

さらに、我々ははじめにでチャンス発見は社会の問題や社会の洞察力の予兆発見が研究対象であり我々は歴史上の社会の問題や社会の洞察力の指し示す方向が研究対象である事をのべた。したがって、この両者を組み合わせれば、将来において社会の問題となる前に予兆の段階で新たに向かうべき方向を探る可能性はあると思う。

謝辞

独立行政法人 産業技術総合研究所の水野光一元技術情報部門長には逆読みがありうる事を指摘いただいた。築根秀男審議役には近くで見守っていただき、おりにふれ励まして頂いた。富士原寛技術情報部門長には日頃から適切なお助言を頂いた。黒澤富蔵元技術情報室長にはこの論文を作成するに当って貴重なお助言を頂いた。高橋千春氏、内藤耕氏、大井健太氏、関根重幸氏、宮本耕一氏始め技術情報部門内の皆様には何度か議論をお願いした。心から感謝します。

参考文献

- 1) 大澤幸生(編著)“チャンス発見の情報技術”東京電機大学出版局、2003
- 2) 博報堂フォーサイトチーム“亜州未来図2010” 阪急コミュニケーションズ
- 3) 庄司裕子、堀光浩一「オンラインショッピングシステムのインタフェースの向上へ向けて」情報処理学会論文誌、Vol.42, No.6, pp.1387-1400, 2001
- 4) 小笠原敦“日本の新しいForesight Program” AIST TECHNO INFO 第23号
- 5) 安平哲太郎著 “概念分析” 文芸社
- 6) 吉川弘之著 “科学者の新しい役割” 岩波書店
- 7) 文部科学省重点領域研究「高度技術社会のパースペクティブ」(代表 竹内 啓 明治学院大学教授)報告書 平成4年9月発行

熊本県医師家塾調明治6年による医療教育と使用教科書 (35事例)

Training Curriculum and Textbook for Medical Doctor on Apprenticeship on 35 Cases in Kumamoto Prefecture 1873.

安澤 秀一：元国文学研究資料館（アーカイブズ研究系）

Dr. Shuichi YASUZAWA, Institute of Japanese Literature
(Division of Historical Archival Study).

yasushuu@coral.ocn.ne.jp

<要旨>明治6年 1873、白川県（明治22年熊本県と改称）は文部省布告に従って、県下の医師で家塾において医療教育を行っている者に対し、本人の住所・姓名・年令、弟子の族籍・姓名・年令、診療する医療の種類について、調査報告を提出させた。35家の報告書には塾則や使用教科書を附属させて医療教育のありようが記入されている。この報告書に記入されている諸データを分析し、明治初年における医療教育の水準を考察し、近代化への道筋において医療の知識普及がどのようになされたか、を考察する試みである。

<Resume>At 1873 DAIJYOKAN (Meiji Cabinet) and Minister of Education, declared to each local governments an investigation about Training for Medical Doctor. SHIRAKAWA (KUMAMOTO) prefecture as local government receive the Report of 35 cases, 12 person in Town of Kumamoto, and 23 person at counties. This article is the result of analysis for the Reports, include status, age, administration rule, fee, medical subjects, and textbooks,

<キーワード>医学教科書, 医療, 漢方, 西洋医学, 製薬, 薬種, 診療科目,

<key ward>medical textbook, medical care, Chinese medical method, western medial method, medicine, drugs, medical treatment,

<本文>

1. 基礎データの性格 慶応4年いわゆる大政奉還によって、大名領はその領国統治を藩という行政単位に移行し、さらに明治4年いわゆる廃藩置県によって、県郡区村の地方行政制度に移行した。近代医学教育制度が確立するまで、医療知識の教育は幕府（万延元年設置）・大名・藩・県の設置する制度的な医学校、もしくは任意の私塾に依存していた。一例を挙げれば佐賀鍋島藩は天保6年以来、藩費で医学寮を運営した。愛媛県は明治7年に医学校と病院を開設し、医学校校長として旧幕府医師、明治5年アメリカ・ニューヨークおよびフィラデルフィア製薬学校に学んだ太田雄寧25歳を招いた。1年の任期終了後、太田は帰京し、医師かつ家塾を開業し、多数の著述と翻訳を行い、さらに明治9年暮「東京医事新誌」刊行を申請し、翌1月から発刊した（太田安雄著『太田雄寧伝』2004刊）。惜しむらく明治14年病を得て31歳の若さで逝去した。医療・医薬についての新しい情報と知識の普及に身命を捧げた先駆者であった。本報告は地方における医療知識形成が、明治初年に、どのように行われたか、を探ろうとするものである。

肥後細川藩にも江戸時代すでに藩医学校があり、明治3年の熊本藩財政資料にも好生館（医局病院）経費が計上されており（熊本県立図書館蔵熊本県公文類纂 32-4 諸藩職制<撮影番号 0015>）、他県のように医学校を併設していたと思われる。

廃藩置県後、熊本藩と八代藩、さらに旧幕領天草が合併して白川県となる。明治6年諸府県に対して、「文部省布告第百一号」（銅版印刷・半紙判小・中折り）が交付された。この調査指示は太政官布告をうけたものである（熊本県立図書館蔵熊本県公文録32-13 医師家塾調明治6年、以下引用の際、出典を省略する）。布告における記載事項を次に掲げる。

布告済其府県管下医師家塾相開キ医術教授スル者ハ別紙雛形ニ照準シ取調早早可差出候、
此段相達候也（手書き）

明治六年七月九日 文部省三等出仕 正五位 田中不二磨

雛形

何府県第何大区何小区何町村何番屋敷居住医師 苗字 名、当何月何歳何月
一塾則、一弟子何人族籍姓名年齢、一弟子何人族籍姓名年齢、一技術何々 内科、外科、
眼科、産科、整骨、綱帯科ヲ実地伝習仕候

右之通相違無御座候以上、 明治六年何月幾日、医師苗字名

「医師家塾調」綴は各医師からの自筆報告書を成冊したもの、美濃版黒色罫紙の提出書類本文には受付担当者による雛形準拠の訂正書込みがある。この種の報告書が熊本県以外の県庁文書にも存在する可能性は高い。ともあれこの調査データは明治初年における医療という健康環境向上への関心の在り様、また知識水準や医療教科書採用の実態、いいかえれば在来医療である漢方、あるいは先進的な西洋医療の教育実態について、愛媛県のように県立医学校・病院を設置する直前、中央政府が全国調査を意図した布告の成果であり、家塾という形態での教育方法の限界と、使用教科書に関する指標となるであろう。

2. 医師と弟子の年齢と出自 家塾運営を届出た医師は35家である。もちろん白川県域内における医師を職業とする数は35家に限らないであろう。ともあれ12家は熊本町居住、23家は郡部各域に広く散在している。家塾医師35人の満年齢は最若年18歳、最高齢62歳、平均年齢は40.1歳である。弟子107人の年齢は最若年14歳、最高齢54歳であり、平均年齢は20.6歳である。ちなみに江戸時代は数え年齢が普通であったのが、この調書で満年齢および月数を記入しているのが注目される。

熊本町居住医師のうち5人が士族である。郡部居住医師のうち士族は1人だけである。弟子は1人から9人までと幅があり、1家塾当たり平均は3人となる。弟子の出自身分は上族医師10人、上族52人、卒医師1人、卒1人、平民医師1人、平民17人、無記入25人、医生1人、計107人である。維新後の新しい状況の中で家禄を離れた士族子弟が医師を目指したようであり、弟子の半数をやや超える。族籍無記入は平民の出自と見てよいであろう。平民もまた新しい専門職を目指したのでであろう。出身地は白川県だけでなく長崎県・福岡県・佐賀県・大分県・鹿児島県・広島県と広がりを見せている。

3. 塾則 塾則は家塾の教育方針と方法がシステムチックに行われているか、どうか、を反映するものと考えられる。塾則記述は繁簡さまざまである。そこで塾則記述の要点をもって類別してみよう。漢方4家、漢方・時間割・製薬1家、漢方・西洋医学1家、時間割のみ1家、時間割・手術・倫理1家、時間割・製薬13家、時間割・製薬・綱帯1家、西洋医学1家、西洋医学・眼科1家、倫理4家、塾則無記述7家という内訳である。35家すべてで実習が行われている。最も詳細な記述を持つ西 道庵26歳・弟子3人（熊本町近在）の塾則（医師番号27）を紹介してみる。塾則が公示されているか、どうかは不明。

1. 日々一員ツツ当番心得之事、2. 診察ハ午前 8 時ヨリ 12 時迄寄留人[入院患者]ヨリ始リ貴賤ノ區別ナク着到ノ順タルベキ事、3. 調薬ハ午後 2 時ニ全ク終ラサル分ハ当番受持タルヘキ事、4. 非番ノ輩午後 4 時ヨリ病家診察ノ事、5. 5・10 之日午後 2 時ヨリ手術療治之事、6. 急病人ハ規則限ニ非サル事、7. 通例製薬ハ当番受持タルヘキ事、8. 寄留願之病人ハ近傍ノ身元タシカナル者或ハ其区戸長之証書ヲ請取ルヘキ事、9. 寄留人ニ関係スル事務并器械等盡ク当番ノ取扱タルヘキ事、10. 寄留人午後 6 時門限ノ事、11. 午後 2 時ヨリ 4 時迄書籍勉学事、12. 禁酒・焼酒ヘキ事

専門分野を明示しているのは熊本町で開業している行徳 文卿 49 歳、弟子 2 人である。「東京教師慘再捏児氏ヨリ伝習ノ方術ヲ主トシ、兼テ抱氏朋氏眼科諸書ヲ採用ス」

行徳 文卿は東京で学んだ西洋医学に加えて、眼科を専門としたようである。なお眼科を診療科目としている者はもう一人居る。明治 6 年という時期ではまだ、西洋医学を主として教育する医師は 8 家であり、弟子数は計 37 人である。

課業の時間配分は家塾によって異なる。午前中診療もしくは製薬調剤、午後後半または夕刻以降に読書と、かなり精励の様子が見られる。読書に関し「行余力有ルトキハ西洋医流ノ翻訳書ヲ学ビ、博ク智識ヲ得ルノ階梯ト為ントス」[医師番号 1]という漢方医の文言をみる。起床は 6 時、就寝は 10 時半という規定[医師番号 16]があり、この家塾では「飲食、家人同様」ともしている。「6 時起褥」[医師番号 7]「午後 8 時ヨリ 11 時迄講義」ともある[医師番号 33]。他塾でもおおむねこうした生活であったであろう。授業料について触れているのは 1 家のみであり、「束修等者勝手次第」[医師番号 23]という。つまり弟子の懐次第で額の多寡は問わない、ということになる。他塾も同様であったであろうか。

塾則に長短さまざまあるものの、塾則を全く記入していない家塾が 6 家、禁酒あるいは博愛衆要術之精といった倫理目標のみが 5 家ある。教育目標と方法がはっきりと見えない場合が三分の一近くになる。家塾運営の制度的整備の限界を示すものといえよう。

4. 診療科目類別 家塾に弟子を置く医師の診療科目を見てみよう。内科のみが 10 人、外科のみは 1 人で、内科・外科を兼ねる者 11 人である。その他、外科の他に整骨と解剖類を加える者 1 人、外科で内科、整骨、綱帯科、眼科まで扱う者 1 人、また外科・内科・眼科、口科とやや趣を殊にする者 1 人、内科・外科・綱帯科 1 人がいる。また内科・外科に眼科を扱うもの 2 人と、内科・外科・眼科に加えて産科・解剖科という者 1 人がいる。甚だしいのは内科・外科・眼科、産科、口中科、解剖、綱帯、整骨、種痘に至るまで扱う者も 1 人いる。また産科を主とするもの 1 人いるが、産科の他に内科・外科を加える者も 1 人いる。例外的に眼科だけの 2 人と、産科ないし小児科だけを扱う者 1 人ずつがいる。医師 35 人のうち 22 人まで内科もしくは内科兼外科と診療科目を限定している。眼科、産科、小児科に専門化している者もいるが、9 人は多岐にわたる診療科目を扱っている。病院と家庭医に分化する以前の状態であろう。診療科目が多い者で弟子 9 人という場合もあるが、内科のみで弟子 8 人という場合もあり、診療科目の多寡によって医療を習得したい弟子を引き付ける誘因となる、というようには見えない。

5. 使用教科書類別 使用されている医学教科書は著者名や刊記なし、書名のみが列挙されている。教科書現物ではないため、手写本なのか、木版本なのか、銅版本なのか、確認できない。書誌として不十分であり、知識普及に媒体がどのように影響するかは不明

6. 医療用薬品 塾則によると、製薬調剤に携る時間がかかなり比重を占めている。漢方であろうと西洋医学であろうと、医療行為として病氣治癒のために適切な医療用薬品を投与しなければならない。家塾調には具体的な医療用薬品の品目が記入されていない。そこで幕末一明治初年に作成された熊本町方値段調帳天保10年一明治3年（永青文庫蔵熊本大学図書館寄託）に依拠し、薬種業者が扱っている医薬品の品目および単位当り価格と対米価指数について、明治2年のデータを列挙してみよう。対米価指数とは明治2年1869米価白米1升＝銀16匁を基準とした相対指数である。

唐物：大黃1斤極上代銀40匁250.0
 甘草1斤大筋代銀35匁218.8
 麻黄1斤代銀15匁93.8
 桂木1斤代銀30匁187.5
 石膏1斤代銀4匁25.0
 大風子1斤代銀6.5匁40.6
 檳榔子1斤代銀7匁43.8

木香上1斤代銀70匁437.5
 山販来極上1斤代銀50匁312.5
 黄芩1斤代銀22匁137.5
 連翹1斤代銀35匁218.8
 蒼朮1斤代銀15匁93.8
 黄耆1斤代銀45匁281.3

和物：半夏1斤代銀30匁187.5
 川芎1斤代銀10匁81.3
 黄連1斤代銀90匁562.5
 当販1斤代銀12匁75.0
 柴胡1斤代銀12.5匁78.1
 澤瀉1斤代銀32.5匁203.1
 陳皮1斤代銀2匁20.0
 姜活1斤代銀7匁43.8
 芍薬1斤代銀13匁130.0
 葛根1斤代銀3匁5分21.9
 大棗1斤代銀25匁156.3
 杏仁1斤代銀70匁437.5
 木通1斤代銀3匁5分21.9
 防風1斤代銀7匁5分46.9
 薄荷1斤代銀10匁62.5
 香附子1斤代銀3匁18.8
 五味子1斤代銀25匁156.3
 山梔子1斤代銀15匁93.8
 厚朴1斤代銀2匁12.5

桂根1斤代銀6匁5分40.6
 天瓜粉1斤代銀10匁62.5
 白朮1斤代銀12匁5分78.1
 桃仁1斤代銀45匁281.3
 牡丹皮1斤代銀80匁500.0
 荊芥1斤代銀13匁81.3
 桔梗1斤代銀10匁62.5
 麦門冬1斤代銀18匁112.5
 三稜1斤代銀80匁37.5
 蘇木1斤代銀10匁62.5
 芎本1斤代銀6匁5分40.6
 細辛1斤代銀13匁81.3
 竹節人參1斤代銀28匁280.0
 薔薇仁1斤代銀33匁206.3
 烏須1斤代銀7匁43.8
 生生姜1斤代銀10匁62.5
 前胡1斤代銀7匁五分
 地黄1斤代銀18匁12.5
 砂参1斤代銀6匁5分40.6

その他：広東人參1両代銀10匁625.0
 御種人參1両代銀35匁218.8
 竜腦1両代銀50匁312.5
 辰砂大1両代銀35匁218.8
 キナキナ1両代銀12.5匁78.1

マクネシア1両代銀4匁25.0
 朱1両代銀10匁62.5
 水銀1両代銀8匁8匁50.0
 烏犀角1分代銀7匁43.8
 麝香極上1分代銀70匁437.5

猪膽 1 分代銀 15 匁 93.8

熊膽 1 分代銀 60 匁 375.0

一角 1 分代銀 3 匁 5 分 21.9

サフラン 1 分代銀 2 匁 5 分 15.6

キナソート 1 分代銀 5 匁 31.3

セメンソート 1 分代銀 4 匁 25.0

薬品の中に、在来漢方薬のほか、片仮名書き 5 品目のあることに注目されたい。

なお佐賀県旧藩主鍋島家伝来・佐賀県立図書館寄託「鍋島文庫」に「製薬品代金記精鍊所」という標題で、明治 9 年 11 月・12 年 4 月・5 月・6 月の 4 ヶ月分を合冊した資料がある。この記録は明治初年の技術水準を示すデータとして興味深いものがある。ここでは参考データとして、明治 9 年 11 月の製薬品代金記における 70 品目の売上記録のみを紹介してみよう。品目名と百匁当り価格（単位円）を列挙する。

硫酸 1.721・硝酸 0.802・塩酸 0.401・醋酸 2.674・丁子油 2.540・炭酸銅 0.909・
 橙皮丁幾 0.104・大黃エキス 0.241・橙皮油 4.543・桂皮水 0.040・剥屋福烏多脂
 1.337・瀘銅卷 3.431・硫化加里 0.368・第二沃求 2.674・龍動鉄丁幾 1.148・ホフマ
 ン鎮痛液 1.337・龍腦丁幾 1.070・格倫模丁幾 0.401・没食酸 70.588・青酸加里
 1.146・沃陣 9.090・礪砂精 1.067・民蛭里精 0.891・鉄雜弗蘭 0.549・ラービス
 20.000・チキタリス丁幾 0.267・神経瀉 1.089・収斂垂的 ■ 1.584・硫酸亜鉛
 0.400・酸化鉛 0.114・鉛糠 1.039・金硫黄 0.800・金液 8.000・薬用石鹼 0.668・洗
 用石鹼 0.230・湿道耳永 1.003・来硝石精 1.069・薄荷油 1.337・小蘭香油 2.286・
 没食丁幾 0.100・琉魂蜜 0.180・硫酸亜的兒 3.075・蔭草丁幾 0.191・猛水 0.923・
 霸王塩 53.833・ヘレヒタート 2.000・燐 19.114・炭酸曹達 0.134・廉働湿湿亜
 0.201・礪砂揮発華 0.228・瀉利塩（斤） 4.010・コーセニール 1.137・アンタリコ
 カリ 0.219・忽布 6.350・龍口草エキス 0.100・海葱蜜 0.200・（製糖蜜（斤）
 10.100・阿窺丁幾 0.214・エレキシールカリサヤル 1.110・鉄鐵塩 0.753・四十度ア
 ルコール 0.735・三十五度アルコール 0.428・三十度アルコール 0.312・苛性加里油
 0.200・精製鹿角精 0.936・精製硝石 0.036・ホッターズ 0.267・

7. 終わりに 明治 3 年熊本藩戸口は士族 4,281 戸、男女計 16,050 人、卒族 15,118 戸、男女計 71,733 人、平民族 120,475 戸、594,321 人、その他合計 149,193 戸、719,990 人であった。こうした人口規模に対する医療従事者を増加させる方策立案のための調査が「医師家塾調」であったに違いない。明治新政府は文明開化・興業自由を目指した。人民の健康衛生を支える医薬業従事の人材養成を旨として、病院と医療教育施設の充実を図った。地方行政を担う藩、そして廃藩置県後の縣において病院・医学校設置が見られた。本報告はその一端として熊本県における民間医師育成を取り上げ、また医薬品供給の状態を瞥見してみた。廃藩置県後の熊本県庁文書の中に病院運営および設備・備品の予算が存在する。病院予算事例は島根県庁文書にも存在する。近代化への転換過程における医療制度の実態解明に資するものといえよう。

謝辞：資料閲覧について熊本県立図書館、そして当時の担当係長平洋子さん始め、お世話下さった皆さんにお礼申し上げます。また熊本大学図書館寄託永青文庫利用については財団法人永青文庫、佐賀精鍊所資料については鍋島報効会のご好意によることを記して感謝する。本データ収集は 2002 年度江戸モノ作り科研助成および佐賀幕末科学技術史研究会の支援による。

初代会長・米田幸夫先生ご逝去

追悼の辞

情報知識学会の初代会長東京大学名誉教授米田幸夫先生が、平成 17 年 11 月 25 日午前 9 時 2 分に肺気腫でご逝去されました。享年 83 歳でした。

米田先生は、大正 11 年 3 月 12 日兵庫県に生まれ、昭和 19 年 9 月東京帝国大学第一工学部応用化学科をご卒業後、昭和 36 年 7 月に東京大学工学部教授（合成化学科）に任せ



られ、昭和 57 年 4 月退職され同年 5 月には東京大学名誉教授の称号を授与されました。

先生は触媒化学の分野で種々の先導的な役割を果たしてこられました。不均一系触媒反応の特に複雑で多様性に富む触媒の研究には、大量のデータ・情報の体系的、定量的な処理の重要性の観点から、計算機の積極的活用の必要性を予測され、その面でも独創的な研究をされました。そのために、先生が深く係わっておられた、有機化合物、有機金属化合物の属性を各種数値情報から推算するシステムである EROICA は、我が国で生まれた独創的な情報化学の成果として、国際的にも高く評価されております。

さらに先生は日本学術会議の学術データ研連（国際データ協議会：CODATA 対応）の指導者のお一人として国内はもとより国際的に活躍され、また情報・知識などを対象とする体系的な「情報学」「情報知識学」の振興の必要性を強く訴えられ、新しい情報学研究連絡委員会（国際情報ドキュメンテーション連盟：FID 対応）の設立にも御尽力いただき広く情報知識学の発展と産官学の連携による共同研究体制にも努力され、材料・反応設計などの自己組織化システムなどの振興にも指導的役割を果たされました。本学会の礎は組織化の面でも具体的な研究指導の面でも米田先生に負うところ甚大と申しても過言ではありません。

独創的な研究とアルコールをこよなく愛された、米田先生のご逝去は誠に痛惜の念に耐えません。ここに謹んで心からの哀悼の意を表し、先生のご冥福をお祈り申し上げますとともに、先生の拓かれた道を辿りつつあるものを温かく見守っていただくようお願い申し上げます。

「情報知識学会誌」投稿規定

2002年8月27日 制定

2003年3月19日 一部改訂

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

1. 投稿資格

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

2. 投稿原稿

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

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

(1) 研究論文 (Research Paper) : オリジナルな研究論文で、内容の主要な部分が学術論文として他に公表されていないもの。

(2) 事例／調査報告 (Report) : 情報知識学に関連したシステムなどの開発、利用、調査に関するもの。資料も含む。

(3) 解説／展望 (Review) : 情報知識に関連した特定分野の論文や学説などを総括、解説、紹介、あるいは技術動向などを展望したもの。技術、研究上の処理、解析方法などに関する解説。

(4) 論談 (Proposal Paper) : 情報知識学に関連した新たな意見の表明、提案など。

(5) 討論 (Discussion) : 本会誌に掲載された論文についての学術的な討論。

(6) 研究速報 (Notes) : 技術、手法、新事実などの簡単な報告。

(7) 講座 (Lecture) : 情報知識学の各分野に関する基礎理論、技術の適用などについて、テーマを定めて系統的に説明するもの。

(8) 学会記事 (News) : 本会の事業、運営などの報告、記事、資料など。

(9) ニュース、お知らせ (News) : ニュース、お知らせ。最近刊行された単行本やモノグラフの紹介。

(10) 講演 (Lecture) : 特別号などにおける講演資料。

(11) その他 : 編集委員会が適当と判断したもの。

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

3. 投稿原稿

3.1 原稿の形式

(1) 投稿時の原稿

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

- A 刷り上り原稿を想定したレイアウト (A4判, 2段組, 20字×46行×2段).
図, 表は希望の位置に配置すること.
- B ベタ打ち原稿 (A4判, 40文字×40行).
図, 表は, 1枚ずつ別の用紙に印刷すること. ベタ打ち原稿右余白に図表の挿入位置を朱書きすること.
- (2) 採択決定後の原稿
以下の C, D の両方の形式で記録された電子媒体 1 部の提出とする.
原稿の送付にあたってはフロッピーディスクなどの適切な電子媒体とする. 詳細は提出時に事務局に相談のこと.
- C Microsoft Word, 一太郎, DVI, PDF などの代表的なフォーマット.
- D 図, 表は十分な品質で印刷できる形式 (JPEG, GIF など).

3.2 原稿の制限

- (1) 原稿の長さを原則として次のように制限する.
研究論文, 事例/調査報告, 解説/展望, 論談 : 刷り上がり 20 ページ以内
討論, 研究速報, 講座 : 刷り上がり 6 ページ以内
ニュース他 : 刷り上がり 2 ページ以内
- (2) 図原稿 (原図) の大きさは A3 判を越えないものとする.
- (3) 原則として, 図版も含めてモノクロ印刷とする. ただし, カラーでなければならぬ図版を使用する場合は, 別途編集委員会と相談する. なお, カラーページやページを超過する分については, 印刷費を著者の全額負担とする.
- (4) 使用言語は日本語または英語とする.

4. 原稿の採否

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

5. 査読のプロセス

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

6. 校正のプロセス

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

7. 別刷

別刷 (抜刷) は著者の実費負担とする. 希望部数を事務局に申し出ること.

8. 投稿の手続き

原稿投稿時には下記の書類を添え, 原稿送付先に郵送する.

8.1 必要書類

- (1) 最初の投稿時 (a, b, c の 3 つが必要です.)
 - 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

- ・論文種別, 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

- 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.
- 3.1(2) に指定した電子媒体: 1部.
- 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: jsik@nifty.com

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.
- 10.3 本学会誌に掲載された執筆内容が第三者の著作権を侵害するなどの指摘がなされた場合には, 執筆者がその責任を負う.

11. 規定の改訂

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

12. 施行

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

13. 改訂履歴

2003年3月19日 一部改訂. 「10. 著作権」に, 10.3項を追加.

「情報知識学会誌」執筆要領

2002年8月27日 制定

2003年5月2日 一部改訂

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200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下部に，表は上部とする．

(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. 章番号は持ちない. 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 (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 200 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。

8. 改訂履歴

2003 年 5 月 2 日一部改訂。英語要旨の長さを 500 語から 200 語に変更。図、表のキャプション位置を訂正。

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 (投稿規定の2.2からお選びください。)
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ぬ場合など, 全額実費著者負担)
10. 投稿日

11. 登録番号
12. 受付日 (再受付日)
13. 受理日
14. 送付掲載原稿
 フロッピーディスクなど 枚数
 プレインテキストプリント 枚数 (刷り上がり見本, 図表なども貼り込んだもの)
15. 別刷り (抜刷り) の希望部数 (全額実費負担)

【投稿原稿整理カード】

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

【掲載原稿整理カード】

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

事務局からのお知らせ

[1] 平成18年度年会費の納入をお願いします

平成18年度(2006年4月1日～2007年3月末日)の年会費を、5月末日までに郵便局または銀行の下記口座へお振込願います。1年分の年会費は正会員8千円、学生会員4千円です。

前年度分未納のかたは合計額を納入してください。請求書が必要なかたは、その旨、事務局へお知らせ願います。

1. 振込先 (振込手数料はご本人負担をお願いします)

- a. 郵便振替口座 00150-8-706543 情報知識学会 (代表 細野公男)
- b. 三菱東京UFJ銀行 秋葉原駅前支店 普通預金 3586133 情報知識学会 (会長 細野公男)

2. 納入した年月日の確認方法

情報知識学会から郵送された封筒の宛名ラベルをご覧ください。[]内に過去4年間、ご自分の納入日が印字されているので確認できます。納入年(西暦の下2桁)、月(2桁)、日(2桁)の6桁です。年会費を滞納している場合は、[未納]と表示してあります。金融機関へ振り込まれてから事務局へ通知が届き、宛名ラベルに印字、発送するまで10日ほどかかりますので、ご了承ください。

[2] 最近1ヶ月以内に事務局からのメールを受信しなかったかた

現在、8割以上の会員がメールアドレスを事務局へ登録されています。各部会の活動や月例懇話会の予告・報告など、頻繁に受発信しており、電子メールは必須の連絡手段となりました。過去に登録されても、最近1ヶ月以内に情報知識学会事務局からメールを1通も受信しなかったかたは、不達が予想されますので、再度、アドレスを事務局jsik@nifty.comへご連絡ください。添付ファイルが開けないかたも、お知らせくださればテキスト文に直して送信します。

[3] 電話でのお問い合わせ

事務局の業務は土日祝日を除き、月曜から金曜日までの毎日行っています。お問い合わせなどの電話は、できるだけ午後1時半から5時までをお願いします。連絡には電子メールやFAXも、どうぞご利用ください。

入会ご希望のかたには入会申込書を、郵送またはFAX送信でお届けします。

情報知識学会事務局

〒110-8560 東京都台東区台東1-5凸版印刷(株)内

TEL:03-3835-5692 FAX:03-3837-0368

E-mail:jsik@nifty.com URL:http://www.jsik.jp

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

芦野俊宏	東洋大学	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	岩田覚	東京大学
内田努	北海道大学	岡本由起子	東京家政学院大人文学部
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	長田孝治	(株)システムソフト
中山伸一	筑波大学	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	藤井賢一	産業技術総合研究所
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
山本昭	愛知大学文学部	山本毅雄	国立情報学研究所

■複写をされる方に

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

〒107-0052 東京都港区赤坂9-6-41 乃木坂ビル 学術著作権協会

TEL: 03-3475-5618 FAX: 03-3475-5619 E-mail: naka-atsu@muji.biglobe.ne.jp

アメリカ合衆国における複写については、次に連絡してください。

Copyright Clearance Center, Inc. 222 Rosewood Drive, Danvers, MA. 01923, USA

TEL: 978-750-8400 FAX: 978-750-4744 URL: <http://www.copyright.com/>

情報知識学会誌 Vol.16, No.2 2006年5月26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 細野公男

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: jsik@nifty.com

URL: <http://www.jsik.jp/>

Journal of Japan Society of Information and Knowledge

~~~~~ Contents ~~~~~

|                                                                                                                                                   |                                                                            |
|---------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|
| A Pattern analysis of award of Grant-in-Aid for Scientific Research with Research Outputs .....                                                   |                                                                            |
| Masaki NISHIZAWA, Masamitsu NEGISHI, Morio SHIBAYAMA, Yuan SUN,<br>Hiroyasu NOMURA, Yoshitaka MITSUDA, Masashi MAEDA ...                          | 1                                                                          |
| Research Linkage between University and Industry in Japan --- Comparison based on NCR-J and<br>CJP databases .....                                | Yuan SUN, Masamitsu NEGISHI, Masaki NISHIZAWA, Keiko WATANABE.....7        |
| Formulating methods of university rankings by publication and citation statistics in ISI's<br>citation databases, and some related problems ..... | Masamitsu NEGISHI ...13                                                    |
| Assigning kansei keywords automatically based on the optimal keywords which experts assigned<br>.....                                             | Takashi HARADA...19                                                        |
| A Method of Reconstruction of Digital Contents .....                                                                                              |                                                                            |
| .....                                                                                                                                             | Norihiko UDA, Atsushi MATSUMURA and Noriko MURATA ... 23                   |
| Features of information seeking behaviors on the WWW for problem solving ...                                                                      | Yoshihiro SAGARA...29                                                      |
| Quantified Estimation Method of User's Profile based on Search Query.....                                                                         |                                                                            |
| .....                                                                                                                                             | Yukio HORI, Takashi NAKAYAMA ...33                                         |
| The Construction of Web-version Information Retrieval System on the Database of<br>"Housei Keizai Shakai Ronbun Souran" .....                     | Hiroki FURUZUMI, Setsuo SUOH ...39                                         |
| Use of a large-scale dictionary on a scientific and technological document retrieval system.....                                                  |                                                                            |
| .....                                                                                                                                             | Akira KODA, Utako MORITA ... 45                                            |
| Examination of application of Structured Model for Museum Object Information ...                                                                  | Nobuo SAITO ...49                                                          |
| Information Processing of Brain and Problem Solving .....                                                                                         | Masao FUKUNAGA ... 53                                                      |
| Healthcare Supporting Systems over the Internet .....                                                                                             |                                                                            |
| .....                                                                                                                                             | Hiroaki SATOMURA, Takehiko TANAKA, Takuya YOSHIHIRO, Masaru NAKAGAWA ...59 |
| Keystroke Dynamics in Text Typing .....                                                                                                           | Toshiharu SAMURA and Haruhiko NISHIMURA ...63                              |
| A Study of relations of the architecture and the document .....                                                                                   | Takahiko YAMAMOTO... 69                                                    |
| A view point as a criterion of judgment seen into from our history .....                                                                          |                                                                            |
| .....                                                                                                                                             | Tetsutaro YASUHIRA, Masataka SAWAKI ...75                                  |
| Training Curriculum and Textbook for Medical Doctor on Apprenticeship on 35 Cases in<br>Kumamoto Prefecture 1873. ....                            | Shuichi YASUZAWA ... 81                                                    |

## Information

|                                   |    |
|-----------------------------------|----|
| Obituary: Prof. Yukio YONEDA..... | 87 |
| Information for Authors .....     | 88 |

情報知識学会誌 第16巻2号 2006年5月26日発行

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

TEL:03(3835)5692 FAX:03(3837)0368 E-mail:jsik@nifty.com

URL:http://www.jsik.jp/

(振替:00150-8-706543)

学術刊行物 ISSN 0917-1436