

INFORMATION AND KNOWLEDGE NEWS

情報知識学会ニュースレター

1995 4.1

31

情報知識学会事務局 発行 〒110 東京都台東区台東1-5-1 (凸版印刷隣内) TEL03(3835)5692 FAX03(3837)0368 ISSN0915 1133

本とコンピュータ

慶應義塾大学文学部助教授 中川 純男

わたしの専門分野は西洋古代・中世の哲学の研究である。コンピュータを使用するのは、主としてテキスト・データ・ベースや文献データ・ベースの利用および作成とワープロとしての利用である。しかしながら、歴史的な文献の処理やギリシア語やラテン語を含むテキスト・データの作成は、現在のコンピュータ・システムが最初からそのようなデータの処理を意図して構成されたものではないから、多少の工夫を必要とするし、今後解決しなければならない問題も含んでいる。パーソナル・コンピュータを利用する中で日頃感じている問題をいくつか記してみたい。

○文字情報の使いやすさ

研究の対象となる文字情報は従来、書物という形で提供されていた。したがって書物を読むという作業が研究者の仕事のかなりの部分を占めていた。情報が文字として与えられている以上、読むという作業は今後とも変わらないであろう。コンピュータの利用は、書物に印刷された文字のほかにコンピュータから出力される文字を読む作業を付け加えた。この二種類の文字データを比較してみると、書物という形が、読むという作業と文字情報とのインターフェイスとして、高度に洗練された形であることにあらためて気づく。字の見やすさという点で書物という形が優れたものであることは、コンピュータから出力される文字データを従来の活字印刷物の体裁に近付けることに、現在どれほどの努力が傾けられているかを見れば、ただちに理解できるであろう。文字情報のインターフェイスとしての書物の利点は見やすさだけではない。たとえば、読み慣れた本なら、特定の箇所を探し、あるいは単語を検索するとき、コンピュータのディスプレイに向かってキーボードから文字を入力するより、はるかに早く目的の箇所を見つけることができる。よく読まれる古典の有名な箇所であれば標準的な書物のページを指示されるだけで、どのような内容の箇所であるかがただちに理解される。

(次頁へ)

目次

本とコンピュータ	1	近年の出版流通とコンピュータの利用	8
文科系学部の情報教育における一つの試み	4	インターネットに参加して	12
ターミノロジー・メーリング・リスト		学会カレンダー	14
開設のお知らせ	7	研究報告会・総会開催のお知らせ	15

より使いやすく、より読みやすい形で情報を提供するために、書物は数百年、手書き写本の時代も含めるなら二千年以上の歴史の中で、さまざまな工夫を産み出した。文字の形、レイアウト、さらには単語の分かち書きやハイフネーション、文の区切り、段落、章分け、ページづけ、目次、索引なども書物をより読みやすくするための工夫の産物である。書物にはそこに記された文字の提供する情報の他に、情報を正確に伝えるための工夫、の書物を使いやすくするための工夫が含まれているのである。文字情報をテキスト・データとしてコンピュータ可読な形に置き換えるときも、このような工夫をまったく無視してしまうことができない。コンピュータに蓄えられた情報も、どこかで人間が読むことのできる形に置き換えなければ、利用できないからである。ところが書物においてさまざまな視覚的形態として実現していた工夫は、コンピュータの処理するデータにおいてはすべて、定義によってのみテキストの内容から識別されるコードとして記録される。このことが、従来書物という形で提供されていた情報をコンピュータで利用する上でのさまざまな問題を産み出す原因の一つとなっている。

テキスト・データを使いやすくするためには、書物に似た体裁でそのデータを引き出せることが望ましい。そのためには本来書物に固有な、行、ページ、ハイフネーションといった概念をテキスト・データに持ち込まなければならぬ。ときにフォント、文字の大きさについての情報も必要になる。ところがテキストの内容に関する処理をするとき、これらの情報は無視されなければならない。単語を検索するときには、ハイフネーションによって二行に分けられた単語も検索対象に含まなければならない。したがって内容に関わる情報と、それ以外の情報とは識別可能な仕方でテキスト・データに収められていなければならない。じつ、テキスト・データの多くは既存の書物をベースにしたものであるから、書物に固有のこれらの情報を何らかの仕方で含んでいる。問題はその仕方にさまざまな方式であることである。複数のデータを処理するときには、そのデータに合わせてそのつど設定を変更しなければならない。現在フロッピー、CD-ROMなどの形で提供されているテキストはかなりの数があるが、どのような形でデータが作成されているかは、必ずしも公表されておらず、それぞれ独自の表示、検索プログラムを付属されることで対応している。したがってこれらのデータを利用する場合は、それぞれの検索・表示プログラムを利用する他はない。図書館など、不特定の人が一つの機械で数多くのテキスト・データを扱う場合には、所蔵するデータすべての検索・表示プログラムを予めインストールしておかなければならない。このことがテキスト・データの蓄積に大きな障害となっている。現在CD-ROMで提供されるテキストの数は急速に増えつつあるが、そのデータの数だけの表示・検索ソフトがある現状は、快適な利用環境と言うにはほど遠い。

○テキストの性格に応じた柔軟な規格

規格の統一が急がれることは言うまでもないが、その規格はテキストの性格に即した柔軟なものでなければならないであろう。ここでテキストの性格というとき、大きく分けて二つの問題を念頭に置いている。一つは、現在まだ多くの課題を抱えている多言語処理の問題である。書かれた文字を読みやすくする工夫は言語によって異なっているから、すべての言語に共通の一義的な規格は不可能であろう。この問題についてはしかし、あらためて考えることにしよう。もう一つはテキストの用いられ方の相違である。辞書のような書物と、それ自体が研究の対象とされる書物とは使いやすさのために必要な情報も異なっている。アルファベットを用いて表記される言語の場合、単語の分かち書きとか、文章の区切りは内容の理解にもっとも密接に関係した情報である。これに対し、行の区切りは内容の理解に、直接は関係しないから、テキスト・データの中に含めず、ディスプレイへの表示、あるいは印刷などの際、利用する機器の性能や利用者の好みに応じて変えることができるようにすることが考えられる。ディスプレイの片隅に辞書を表示したいときなどには有効な方法であろう。しかし、すべての書物をこの規格で統一することはできないであろう。辞書のよ

うな使われ方をする書物と、繰り返し読まれ、それ自体が研究の対象とされるような書物とを同列に扱うことはできない。繰り返し開かれる書物の場合、われわれはその内容とともに、ページ全体の視覚的印象なども記憶しており、このことが使いやすさの一因となっているからである。アリストテレスの『形而上学』第一巻980a25（980ページa欄25行目）といった箇所の指定方法の代わりにファイルMET_1.TXTの297バイト目から330バイト目までといった指定方法が広く受け入れられるようになるとは信じがたい。

○テキストの確定と共有

文字情報の伝達手段として、活字印刷の書物が可能にした利点は、情報の確定と共有である。この内、情報の確定ということに関して、機械可読データの長所でもある柔軟な改変可能性は、かえって弱点となる。テキストのある箇所が書き換えられたとしても、書き換えられていることはデータそのものからは知られない。これと類似の問題は実は過去にも存在した。活字による出版が行なわれる以前、書物は手書きの写本という形で流通していた。手書きであるから書き誤りや脱落が生じやすい。人間の手による複写物である限り完全に同一のものは期待できない。欄外の注記が本文に紛れ込むこともある。このような危険は活字印刷が行なわれるようになって大幅に減少した。複数の読者が完全に同一のテキストを手にするようになった。あるテキストを研究対象とするとき、そのテキストが確定されていることはきわめて重要である。現在提供されているテキスト・データが何らかの印刷された書物をベースとせざるをえない理由の一端もここにある。CD-ROMのように書き換え困難な媒体を用いて情報を提供するとしても、電子出版の場合、情報が媒体に依存する度合いは活字印刷物に比べればはるかに低いから、確定された情報を保証する何らかの手だてが必要であろう。

古典のテキストの場合、テキストを確定すること自体が実はすでに、知識の集約である。大文字だけが単語の区切りなく記されている写本においては、一つの箇所の単語の区切りについて複数の可能性のあることがある。単語を分ち書きするときには、その複数の可能性から一つを選択しなければならない。あるいは手書きの文字では、ΠとΓとTはしばしば識別困難である。しかし印刷物として公刊するときは、もっとも妥当と思われる一つを決定しなければならない。さらに、多くの古典的な著作の場合は、複数の写本が現存し、そこにはさまざまな差異が認められるから、いずれがより真正なテキストであるかの判断も要求される。このような複数の判断を経て、確定されたテキストができあがる。この意味で公刊されたテキストはそれ自体が知識の集約である。ところが書物という形は、知識の成果を伝えるには適しているが、知識集約のプロセスを伝えるには適していない。

問題がことばの意味にかかわる場合は、さらにたいへんである。ある箇所、ある問題についての複数の解釈、反論、さらに反論への反論などが、さまざまな学術論文として公表される。別の問題について書かれた文献が、解決の糸口を提供することもある。このようなさまざまな文献を、結びつけ、比較することはそれ自体知識の集積であるが、この知識の集積を書物という形に反映することは、容易ではない。活字印刷の書物はテキストの共有を可能にしたが、しかしこの場合の共有は同じ内容の書物を、それぞれの研究者が所有しているという意味での共有にすぎない。テキストについての研究の成果は、それがふたたび活字印刷という形をとらなければ共有の知識とはならないし、それぞれの印刷物は物理的には独立しているから、出版物相互の参照はindexを用いた参照にとどまらざるをえない。しかしコンピュータを利用するなら、関連情報のindexからただちにその内容呼び出すことも、不可能ではない。このようなデータベースのリンクはそれ自体が知識の反映として実現することであるから、複数の研究者が共通のデータベースの構築に参画し、研究の成果を反映できるようになるなら、そのことの持つ意味はきわめて大きい。

文科系学部の情報教育における一つの試み 人工知能をテーマとして

駿河台大学法部 星川 照

．はじめに

最近の先進諸国における情報化の勢いは目を見張るものがある。大会社の本支店がオンラインで結ばれているのはあたりまえだし、小さな町工場でもコンピュータの2つ3つは置いてある。個人でもコンピュータを利用して仕事をこなしている人や、またパソコン通信を楽しんでいる人も多い。つい最近アメリカのゴア副大統領が情報スーパーハイウェイの建設を提唱したことは記憶に新しいし、遅れてならじと我が国でも通産省が同様の計画を打ち上げたりしている。書店の店先には情報スーパーハイウェイのひな型といわれるインターネット関連の書物が山積みされていて、いかに多くの人がこうした問題に関心を持っているかがうかがわれる。まさに、「情報を制するもの世界を制する」である。

さて、こうした状況を反映して、教育の現場でも情報についての教育の必要性が叫ばれている。多くの大学でコンピュータがすでに導入されているし、小学校、中学校そして高等学校にもコンピュータが本格的に導入され始めている。それら教育の現場に持ち込まれた多くのコンピュータはどのように利用されているのだろうか。小中高の各学校における情報教育が本格的に立ち上がるには今少し時間がかかると思われるので、この小論では話を大学に限るとしよう。

．大学における情報教育

一口に情報教育といってもその意味するところはかなり曖昧である。「情報」という言葉自体、様々な場面で様々な意味を持つ。新聞、ラジオ、テレビのニュースは人間の社会的な活動についての「情報」を伝えているし、絵画、音楽などの芸術作品は、言葉を通してではないが、制作者が他者へ伝達したいと思っている「情報」そのものであろう。また、自然の景観

もそれを求めている人がいる限り「情報」である。コンピュータを利用する情報処理という点から見れば、デジタル化され得る「情報」のみが意味のある情報である。なお、「情報」という言葉の意味については高山〔1〕が詳細に分析している。また、情報化社会における「情報」の質の変化については安澤〔2〕を参照されたい。

以上のようなことを踏まえれば、情報教育とは、広い意味では、「情報」とは何かを教育することであろうし、非常に狭い意味では、「コンピュータの使い方」についての教育と言ってよいであろう。後者は最近「コンピュータ・リテラシー」と呼ばれることもある。ただし、「コンピュータ・リテラシー」と言ったときは、コンピュータを使う上での基礎的な素養を意味するが、「コンピュータの使い方」といったときは必ずしもそうではない。例えば、理科系の学部などではコンピュータの能力をぎりぎりまで使わなければならないような場合もあり、そのような「コンピュータの使い方」は基礎的ではなく、高度な使い方であるから、教育も高度に専門的にならざるを得ないであろう。いずれにしても、現在多くの大学で行われている情報教育は、いま述べたような意味での情報教育としては、狭い意味でのそれにより多くの比重がかけられていると思われる。すなわち、情報教育の入門編では、もっぱらワープロや表計算ソフトの使い方を教えるし、応用編では、理科系ではFORTRANやPASCALといったコンピュータ言語を、また文科系、特に経済系などではSPSSやSASといった統計パッケージの使い方などを教えることが多いのである。これらはいずれも「コンピュータの使い方」の教育である。

．一つの試み

筆者の所属する大学には、法学部、経済学

部、文化情報学部3学部の3学部があるが、そのうち法経の両学部では、2年次に教養演習という科目が設置されている。それぞれの学部学生はそれぞれの専門分野の科目を数多く履修することになるが、そのなかにあつて、少しでも多角的な視野を育成しようと設けられたものであつて、名前にも示されているように、ゼミナール形式で行われ、履修学生数も20名以下という少人数である。担当教員は20余名程度であり、演習の内容はそれぞれの担当教員に任されている。少人数教育が約束されており、しかもゼミナール形式ということもあつて、一般の多人数の講義形式の授業に比べて、より大きな教育効果が期待できる。

さて、筆者がこの教養演習を担当することになって、この科目の意義と特徴と自由度を活かすために演習の内容をどのようなものにしたらいかがいろいろ考えたが、その結果、人工知能をテーマとすることにした。その理由は、第一には筆者が興味をもっているテーマということ、第二には文科系学部の学生にとってもおそらく興味のあるテーマであろうということ、第三には同様の内容の講義や演習がないこと、第四には内容から考えてコンピュータによる実習が必須であるが、履修学生数のMAXが20名であるので、教員一人でも十分に対応できること、などである。さらに、このテーマは情報教育の範疇に入るが、工夫次第で、情報教育を単なる「コンピュータの使い方」の教育にとどまらず、もっと幅の広いものにできるのではないかと考えたからでもある。

以下に述べるのは、担当初年度から数年にわたる上記演習の試行錯誤の経過である。

担当初年度は、筆者自身がこのテーマでの演習を経験したことがなかったので、まったくの手探り状態でのスタートであつた。受講生が文科系学生で、コンピュータに触れたことのない学生の方が多い状況の中で、まずキーボードに慣れさせる必要があると思われた。そこで最初の時間はキーボード・トレーニングのソフトを使い、さらに次の数時間はMS-DOSのいくつかのコマンドの練習や、エディターの使い方の練習をした。いよいよ人工知能についての話

となるわけだが、初年度は結局講義形式でやることにした。これではせっかくのゼミナールの利点を十分に生かせないが、そうした理由は、筆者の経験不足から、講義内容に対する学生の反応が予想できなかったためである。すなわち、学生の反応に応じて臨機応変にその後の内容を変化させるのは、教育のテクニックとすればかなり高度なことであり、始めは一方的な講義の方がやりやすいと考えたからである。内容はごく一般的なもので、人工知能の歴史、問題解決システム、自然言語理解・機械翻訳、エキスパートシステム、知的CAIなどである。人工知能についての講義は前期で終了させ、後期はLISPの実習にあてた。LISPはPrologとともに人工知能向き言語の代表と言えるが、PrologではなくLISPを選んだのは筆者の好みからであつて、それ以外に特別な理由はない。LISPには様々な方言があるが、ここではパーソナル・コンピュータ上のLISPとして定評のあるmuLISPを用いた。またmuLISPにはcommon LISPのモジュールが付属していて、それを読み込むことによってほとんどcommon LISPと同様に働く。LISPの実習は、LISPの組み込み関数一つずつ説明して、例題を学生にやらせる、という形をとったが、思ったより多くの学生が落伍することなくつてきた。しかし、一方で何がなんだか判らないという学生も若干名いたし、また、落伍しなかった学生のなかにも、一つ一つの関数の働きは理解できても、それらを組み合わせでプログラムをつくるということがよく理解できない学生もいた。所詮、半期の実習では、プログラムを作るのに十分なことを教えるのはむりだし、まして動かしておもしろいプログラムを自作するのはむりと言える。

初年度の反省を踏まえて、次年度はいくつかの改良を試みた。まず、せっかくのゼミナール形式の演習であるので、テーマに対する主体的な取り組みを期待するという意味で、前年度には講義形式であつた人工知能についての概略を、学生の研究に委ねることにした。具体的には、学生を数名ずつのグループに分け、それぞ

れのグループに研究のテーマも含めて自主的な調査研究を課した。結論から先に言えば、これは失敗であった。一年次に情報概論などの講義を履修した学生が若干名いたとはいえ、多くの学生にとって、いきなり人工知能のことを調べろといっても少し荷が重すぎた。どのグループの発表も、大学生のゼミ発表としてはあまりに幼稚であるか、あるいは参考図書の棒読みにすぎなかった。もう一つの改良は、演習にビジュアルな要素を取り入れたことである。DEC社の制作した「人工知能入門」というビデオ教材で、主にエキスパートシステムに焦点をおいてナレッジ・エンジニア育成を目的としたものである。この試みは失敗とは言えないまでも、成功とは言えなかった。それは、このビデオが人工知能の中でもエキスパートシステムに強く偏っているので、人工知能一般の入門用には適していないからである。

その後、部分的な改良を加えて数年経過したが、昨年よりかなり大きな変更を加えた。それはエキスパートシステム構築用ツールを導入したことである。それまでの数年間、いかにして学生に自発的に人工知能に取り組ませるか、という点に苦心していろいろな方法を試みたが、どれも十分な成果を見だせなかった。エキスパートシステム構築用ツールにしても大きな期待を持っていたわけではない。ところが、学生にエキスパートシステム構築用ツールの使い方の初歩を教え、自分たちでエキスパートシステムを構築せよ、という課題を与えると、積極的に課題に取り組み始めた。学生が考えたテーマをいくつか紹介すると、「カクテル・エキスパートシステム」、「天気予報エキスパートシステム」、「猿の種類エキスパートシステム」などである。中には100を越えるルールをもつエキスパートシステムを構築したグループもあった。どれも学生たちの独創であり、教師からの誘導は一切ない。エキスパートシステムとしては、予算の制約や機器の制約から「創玄」および「大創玄」を併用した。「大創玄」はプロダクション・システムとしては後ろ向き推論のほか前向き推論にも対応しており、さらにフレームベース・システムも構築できるなど機能

が充実しているが、この程度の演習には「創玄」でも十分と思われる。

いずれにしても、エキスパートシステム構築用ツールの導入は非常な成功であった。本格的な人工知能の研究者を育成しようとするのが目的ではないので、LISPなどの人工知能向き言語をむりして教えるより、このような簡易言語によって、実際に動くエキスパートシステムを自力で作る方が、教育的な見地からも適当で、しかも学生にとっても喜びが大きいのではないだろうか。またDEC社のビデオ「人工知能入門」もそのような目的に最適である。ここ数年は基本的にはこのスタイルでいこうと思っている。

．おわりに

以上述べたことは、人工知能をテーマとして演習を数年間担当した経験談のようなものであるが、学兄諸氏のご批判をいただければ何よりの喜びであるし、またこれから同じような講義や演習を考えておられる方に多少なりともご参考になれば望外の幸せである。

数年前までのAIブームがうそのような昨今である。いまさらエキスパートシステムを取り上げても意味があるかについて、筆者も正直なところ若干躊躇した。この点については、寺野[3]の論文に大いに勇気づけられたことを付け加えておきたい。

参考文献

- [1] 高山正也．” 情報の定義：情報とは I n f o r m a t i o n か I n t e l l i g e n c e か ” ， I N F O R M A T I O N A N D K N O W L E D G E N E W S v o l . 2 5 (1 9 9 4) 6
- [2] 安澤秀一．” 情報化社会における一つの試み－駿河台大学文化情報学部のコンセプト－ ” ， i b i d . v o l . 2 5 (1 9 9 4) 4
- [3] 寺野隆雄．” 知識システムのシステム化技術と標準化の動向 ” ， i b i d . v o l . 2 6 (1 9 9 4) 8

ターミノロジー・メーリング・リスト開設のお知らせ

ターミノロジー（専門用語）関係の情報交換を行うため、メーリング・リストを開設しました。使用は無料で、参加・脱退も自由です。インターネットを使用している方、またインターネットへの接続をサービスしている商用 BBS に加入されている方であれば、どなたでもアクセスできます。公法に抵触しない範囲で自由にご利用ください。交換する情報は、以下の項目です。

- 文献情報（辞書・事典、ハンドブック、規格類、論文、雑誌記事等）
- ソフトウェア情報（電子辞書、辞書検索ソフト、パブリックドメイン辞書等）
- イベント情報（学会、国際会議、セミナー、シンポジウム等）

参加方法は以下のとおりです。

（１）一番最初にメンバー登録を行います。これは `subscribe` コマンドを使用して、以下のようにメーリング・リスト・サーバーにメールを送ることにより行います。Subject: はあってもなくても構いません。コマンドは１行目に記述してください。

例： `% mail majordomo@ijnet.or.jp`
`subscribe terminology@ijnet.or.jp`

たとえば、ニフティサーブから参加するかたは、次のメールを送ります。

TO:INET:majordomo@ijnet.or.jp
SUB:xxxx
`subscribe terminology@ijnet.or.jp`

（２）メンバー登録に成功したら、Welcome to terminology というメッセージが返信されるので、この時点から購読と投稿が可能になります。投稿先は `terminology@ijnet.or.jp` です。なお、脱退のコマンドは `unsubscribe` です。

問い合わせ： 金森國臣

〒132 東京都江戸川区小松川 1-2-2-910

電話 (03)3636-8732 E メール GCE01037@niftyserve.or.jp

突然だが皆さんは、書店に本を注文されたことがあるだろうか。たいがいの書店員は「到着まで10日から2週間位かかりますが、それでもよろしいですか」と聞いてくる。本は情報や知識を得るための媒体であるから、代替がきかない。仕方なく「それで結構です」と答え、リファレンスカードに必要項目を記入する。2週間ほどすると、書店から注文した書籍が到着した旨の連絡が入る。そこで出かけて行って、それを定価で購入する。書店員から「大変お待たせして申し訳ありません」という一言をきくことはめったにない。手に入ればまだ良い方だ。ひどいときは2週間ほどして書店から、「実はこの書籍は出版元で品切れした」という連絡が入る。また小さい書店には「客注お断り」という貼り紙がしてあるところもある。貼り紙こそしていなくても、たいがいは店員が断っている。つまり「お店にあるものは売りましょう。ないものは取り寄せできません」と言うことなのである。

日用雑貨など一般消費財ならば、店頭への品揃えはもちろんのこと、欠品している商品の補充は分刻みで行われていると言っても過言ではない。店頭での欠品は即座に売り損じにつながり、その累積が市場戦略に大きな狂いを及ぼすことさえあり得る重大事態なのだ。情報や物の流れをできるだけ、速やかにするために莫大な投資をしているのは、そのためである。

数年前、花王株式会社の佐川会長（当時）の話を聞いたことがある。花王では店頭の商品揃えが会社の命題と位置づけ、24時間体制の情報ネットワークを構築した。夜間もコンピュータを無人で運転し続け、刻々と入る受注情報をリアルタイムに加工し、梱包出荷までを、文字どおり「一刻を争う」速度で次々とこなしていく。これによって、顧客満足という他社との差別化が実現できるのだそうだ。

もう何年も前のことだから、今ではもっと進んでいることだろう。これが一般消費財の流通の最先端の事例であろう。

それにひきかえ、出版物流の遅さはなんと表現したら良いのだろうか。これは一般的に次の2点に起因していると言われている。

その第一は委託制である。新刊書（文字どおり新しく刊行された書物）や雑誌は、取次会社（出版社と書店の間にある問屋）による見計らい配本によって、発売日に書店店頭へ山積みされる。消費者は、新聞広告などのマス媒体によって得た情報によって、書店で現物を見てそれを購入する（あるいはやめる）。そして、それらは一定期間陳列された後、書店が売れ足が鈍ったと判断したら一定期間内ならば自由に返品することができる。

（出版社は、その返品を無条件に引き受けなければならない）。精算は、配本されたものから返品したものをひいた後のもの、つまり売れた分だけ支払えばよい。これが委託制である。

つまり、書店はある種のスペース産業であり、限られたスペースの中で売れ筋のものを並べ、いかに効率よく売るかということに徹していると言っても過言ではない。出版物の総発行数は年間49億冊、その返品率は、平均して40%。一冊単位の注文品にどうしても時間がかかってしまい、その結果「客注お断り」が発生するのは避けられないことなのかもしれない。それに注文品は、本の間にはさまっている「たんざく」が注文書兼納品書の役割を果たすという全く原始的な方法を採用しているので、物流速度を自ずと落としてしまっている。何が原始的かと言えば、なんとこの「たんざく」には20人の指紋がつくと言われている。注文書である「たんざく」は書店員の手によって起票され、取次会社にわたり、各部門を経て出版社に到着する。ここで

注文書の役割は終わる。ここまで来るのに約一週間かかる。そして出版社の倉庫で初めて現物の書籍と「出会い」、今度は納品書として本と一緒に取次会社へ行く。更に取次会社ではトラックの出発を何日か待たされてやっとの思いで書店に到着する。この帰りがまた一週間である。大量の新刊書や雑誌の配本といういわば大量輸送方法に比べて、この注文品の流れは旧態依然としている。

しかし、それらの物流をより迅速にこなすために、他業界のように設備投資をすればいいじゃないかという意見がある。もっともなことだ。FA化を徹底して推進すれば、流通速度はもっと上がって、それこそ「客注お断り」などということもなくなるかもしれない。

ところがその設備投資をできない理由がある。それは、この出版業界の正味（マージン率）の低さである。これが出版物流に時間がかかることの第2の理由である。一般に書店のマージン率は、実売の20～22%。取次会社は8～10%と言われている。1000円の本を売っても書店には200円しか残らない。これでは人件費や家賃など固定費をまかなうのがやっとなで、設備投資など未来費を捻出する余裕などないというわけだ。つまり書店は本を売ってもそれは儲からない。もし客注を受けても、それを電話で出版社に注文を出しやっとな到着した荷物をほどこき、あれば注文主に電話をして取りに来てもらう。いつ来るともわからない相手のためにしなければならない、そんな客注品を扱うぐらいなら、取次会社からのお仕着せのパターン配本によるものだけ売って、残れば返品すればよい、ということになる。

この二つの原因の他に私はさらに二つをつけ加えようと思う。

そのひとつは、これは冒頭にも述べたが、出版物がその特性として極めて代替性の低い商品であるということである。たとえば、A氏のBに関する書物を手に入れたとき、それはN氏のものであっても、またA氏のCに関する書物であってもいけないのだ。「お客はA氏のB論文が欲しいのだから、多少は時

間がかかっても待ってくれるだろう」というある種の甘えの構造が、出版物流には時間がかかるものという、業界の「常識」を作り出してしまったと言える。

もう一点は、本稿の中心テーマである、情報化の立ち後れである。たとえば、書店店頭で各出版社の在庫状況の把握ができれば、前に述べたような、2週間たってから「実は品切れでした」などという事態はなくなるだろうし、発注にしてもPOS端末から自動発注できるシステムを取り入れれば「たんざく」による発注に比べて、所要時間は大幅に改善されるだろう。さりとてこれも、先ほどの正味の低さに起因して、設備投資が進んでいないというのが現状だ。

情報化のインフラであるコード体系の整備にしても、ISBN (International Standard Book Number) というものはあるにせよ、それが書籍自体に表示されていないなかったり、比較的読みとりが簡単にできるバーコードが、体裁上不釣り合いだとして付記しない出版社も数多くあったりで、その道のりはほど遠いと言わざるを得ない。大手出版社主導による「出版VAN」構築の動きもここへ来て「発展的解消」をしてしまった。現実には一部の出版社と取次会社間でのオンライン受注が行われているのみである。

以上見てきたように4つの要因が複雑にからみあって、出版物流は、とりわけ、客注品の流れが遅いことがわかったが、以下はこれを逆手にとり、自社出版物のクイックデリバリーに早くから取り組み、システム化している婦人之友社の事例を紹介する。

婦人之友社は1903年羽仁吉一・もと子夫妻によって設立された老舗出版社である。雑誌「婦人之友」は、現存する雑誌の中で4番目に古い。出版内容は雑誌「婦人之友」の他に、家事・料理・育児・健康と言った実用書を中心に120点ほどの出版物がある。さて、この婦人之友社は今から9年前の昭和61年から「書籍売掛直送」を開始した。それまでの書籍の販売方法は、マス媒体に広告を打ち、

書店での購入を促す。どうしても直送希望という客に対しては、先に本代と送料を送金してもらい、入金確認後書籍を発送するというものだった。しかし婦人之友社の規模では、とうてい全国の書店にまんべんなく新刊書が配本されるというわけではなく、そもそも、買切制（委託制と違って、注文のあった分だけを書店に配本する制度。従って返品は原則として受け付けない）を採用しているため、マス広告を見て欲しいと思った読者が、確実にそれを手に入れることは難しかったのだ。そこで読者に先に本を送って、後で送金してもらう方法を採用した。そしてこれらの注文情報をコンピュータに蓄積して、商品のジャンル毎に分類し、時宜を得て販売促進のDM（Direct Mail）を出状すると高率での反応があった。

これらをイメージ化したのが図1である。

処理の中心はオフィスコンピュータ（三菱電機MELCOM80）である。まず様々な宣伝媒体によって得た顧客をデータベース化する。データベースはRDB（Relational Data Base）を採用している。これはMELCOM80がRDBを高速に処理する専用プロセッサ（GREO）と4GL（第4世代言語－EDUET）を持っているからである。データベースには顧客の個人情報（住所・氏名・年齢・性別など）と購入明細情報（書名・冊数・送り先・注文頻度など）を登録する（処理①）。個人情報と購入明細情報を別ファイルに登録するのは、個人情報ファイルは他システムの預り金情報とファイルを共有するからである。次に購入明細情報を購買履歴情報に加工する（処理②）。これが、将来の販売促進DMを展開する際の重要な資料となる。購入履歴情報にはいわゆるRFM分析の手法に基づいてその項目が設定されている。具体的にはR（Renency = 最新注文年月日）、F（Frequency = 注文方法別回数）、M（Monetary Value = 累積購買金額）その他ジャンル別購買数、DM応答回数などがある。第3に購入明細情報から送

票出荷明細・請求書などの出力処理を行う（処理③）。クイックデリバリーがこのシステムの命題だから、そこまでの処理を短時間に行うため、①には郵便番号辞書、姓名辞書、③には高速プリンタを配備している。

更にこのシステムの特徴は、これら構築したデータベース（顧客）に対してDMを出状すると、かなりの高率でレスポンスがあるということである（処理④）。近年の実績を見ると、常時13～15%のレスポンスがあり、ものによっては20%を越えるものもあった。出版企画に対して相応の顧客を絞り込み、個人情報と結合させて、高速にデータを抽出する。その結果DM規模はどのくらい、収益予想はこのくらいというめやすを、数分の内に立てることができる。これらをDMの「意志決定」と呼んでいる。これはGREOが搭載されたハードであるが故に実現できることである。更に購買履歴情報をRFM分析し、ABCランク設定を行ってより精度の高い顧客情報へと磨き上げていく（処理⑤）。

ただ、これらの事がコンピュータ専門家にしか扱えないとしたらシステムの効率は半減してしまうだろう。しかしながらMELCOM80に標準搭載されている4GLのEDUETを利用することなどによってEUC（End User Computing）環境を整備し、少なくとも営業に携わるものはみな、データの絞り込みができるようにしている。

ここで具体的にEDUETを使ったDM実施リストの抽出例を見てみよう。まずRDBの機能のひとつである選択（Selection）である。EDUETの「レコード抽出条件」画面でどのような条件でレコードを選択するのかを式を使って指示する。たとえば最新注文日が、今年の1月1日以後で5000円以上購入している人ならば

A05（最新注文年月日の項目ID）>= 950101 AND A06>=5000
という具合である。

次に第2のRDBの機能である射影（Projection）である。DMを発行するために、

データベース中のすべての項目が必要であるとは限らない。その際必要な項目だけを取り出すのが射影である。今回は購買履歴情報の中の、顧客コード、最新注文日、累積購入金額の項目だけが必要であるとすれば、「出力項目指定」画面でその項目IDを指定してやればよい。

最後に特定のキーを参照して他ファイル(DB)を取り込むことができるこの機能を結合(Join)という。

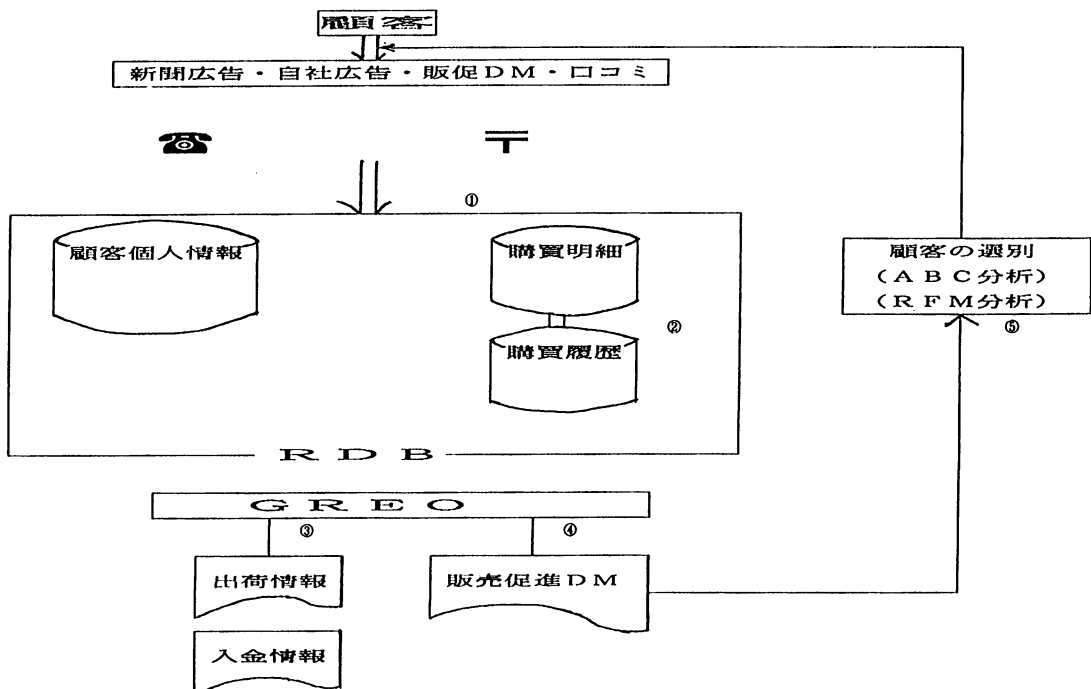
今回はDMのあて名ラベルを作成することにするから、顧客個人情報住所・名前を参照にしなければならない。そこで、「ファイル選択」画面で個人情報ファイルのDB名やキー項目を指定してやればよい。これらの操作はほとんどの人が2日ほどで習得することができ、また、処理速度は10万件のDBの検索で1分以内である。

実用書は、出版物の中でもどちらかと言えば一般消費財に近い存在である。お弁当づくりの本を買おうと思って来店した客は、やはりそれこそよほどのことがない限り、見た目に美しい本、内容が楽しく、簡単につくれる本を購入して行くだろう。その観点から婦人

之友社は、自社出版物を一度手に取った客の情報をデータベース化し、厳選された顧客に対し繰り返しDMを打つことで、ファンを作り、リピーターを開発しているのである。一見簡単ようだが、このDMの高率は長年かかって構築したノウハウの結果であると言える。

この書籍直販システムでは、電話注文の場合最短で中2日で書籍が読者のもとに届く。発送を担当する流通倉庫は遠隔地にあるが、ピッキングリストや出荷明細などはオンラインで伝送している。流通速度において、先に述べた書店の場合とは、格段の相違であることは一目瞭然であろう。このことで他社との差別化を図り、顧客を「固客化」している。フルフィルメント全般についての論は別の機会に譲ることにする。

この事例はすべての出版社に当てはまるとは思えない。それぞれの出版ジャンルや購読対象によって、おのずと戦略は異なるだろう。ただ一つ言えることは、読者は「読みたいときが欲しいとき」なのだ。それをより理想に近い形で具現化したことが、成功のキーにつながると思う。



インターネットに参加して

明海大学外国語学部英米語学科講師 渡辺 雅仁

☆はじめに

巷にはインターネット関連の書籍があふれています。情報ハイウェイ、マルチメディア対応のWWW、ホワイトハウスや首相官邸への接続、などなど、インターネットがバラ色の未来へのキーワードとなっています。インターネットでできることの全体像については、市販の解説書に譲ることとし、ここでは言語学を専門分野とする筆者がどのように活用しているか、事例報告を行いたいと思います。

☆メーリング・リストへの参加

インターネット上では個人の興味に応じてさまざまなフォーラムが形成されています。その一つに「メーリング・リスト(以下「リスト」)」があります。これは、ホストとなるコンピュータに簡単な電子メールを送付するだけで参加することができます。このホストに参加者として登録されると、ホストは参加者全員に、世界中からホストに投稿された記事を自動的に電子メールで送付します。このように送付された電子メールをもとに、特定のテーマで議論したり、情報を交換します。活発なリストでは1日に20~30もの投稿があります。筆者は具体的には、Linguist という言語学と、TESLCA-L という CAI 関連のリストに登録しています。

☆外字による発音記号の限界

筆者は、個人の研究として英単語を発音記号で表示することに取り組んできました。これまでは日本語ワープロ上で「外字」と呼ばれるユーザー定義の文字を使用してきました。ところが、この方法にはいくつかの欠点があります。

1. 外字は漢字と同列に扱われるので、日本語の走らないパソコン上では表示できない。

2. 外字はドットと呼ばれる点によって構成されるので、拡大するときざざが目立つ。

☆TrueType の発音記号

Microsoft Windows はパソコンの世界に大きな変革をもたらしました。その一つは、日本語-英語の壁を壊し、海外で作成されたソフトが日本語の環境で走るようになったことです。これは具体的には、日本語の環境でも海外で作成されたさまざまなフォント(字体)を使用できることを意味します。しかも、このフォントは TrueType と呼ばれ、自由自在に拡大、縮小ができ、大きく拡大してもなめらかな輪郭を維持します。しかし、TrueType は文字をベクトルに分解し、複雑な数式によって文字を記述するため、日本語ワープロ上で外字を作成するように、単純に作成することができません。

☆TESCA-L に Help!

そこで、TESCA-L に「発音記号のフォントセットを誰か知らないか?」とヘルプメッセージをだしました。1週間のうちに4名から、レスポンスがあり、アメリカの SIL(The Summer Institute of Linguistics)社が、IPA(国際音標文字)のフォントセットを発売していることを知りました。次に SIL 社に電子メールを送り、その詳細を尋ねました。すると、次のような電子メールが届きました。

Dear M Watanabe

Thank you for your request for information about SIL fonts. The following is the most current information about the fonts.

SIL Encore Fonts version 2.0:
Microsoft Windows

The SIL Encore fonts are a package of over 1300 phonetic characters and linguistic symbols that can be used to create customized, scaleable fonts in either PostScript Type 1 or TrueType formats. The SIL

Encore glyph library has a Roman-based glyph set for linguistic applications in three font families (SIL Doulos, SIL Sophia & SIL Manuscript) in four type styles (regular, bold, italic and bold italic) available in Macintosh or Windows platforms. A free Windows TrueType font compiler is included with the Windows TrueType package. With Mac or PostScript you would need a third party font manipulation package, such as FontMonger or Fontographer to build your customized font.

The cost of this package is \$100 US (individual license) plus shipping and handling (\$5 US inside the USA and \$20 outside the USA). Site licenses are available. You can order in Windows TrueType or PostScript, Macintosh TrueType or PostScript. Sorry, we do not take credit cards. You can send a check or postal money order for payment.

要約: SIL Encore fonts は 1300 の音声記号から成り立っており、PostScript(プリンター記述言語)版と TrueType 版とがあり、Windows と Mac の両方で使用することができます。価格は 100 ドルで 20 ドルの送料が必要

☆SIL Encore Fonts

国際郵便為替で 120 ドルを SIL 社に送付しこれを購入しました。1300 あるフォントのリストから、必要なものを選択し、キーボードの

タイプフェースにわりあててのみならず、複数のフォントを組み合わせ、オリジナルなフォントを作成することもできました。以下は、筆者が作成した発音記号の一例です。TrueType による美しい仕上がりとは日本語との混在に注目してください。

[æi]	これは11ポイントです
[æbəkəs]	これは18ポイントです
[æbændən]	これは26ポイントです

☆おわりに

今回のことを通じて、世界に散らばった同じ志をもった人たちをつないぐコンピューターネットワークのすばらしさを体験することができました。その基本は、分からないことを質問し、その質問に答えるという人と人とのコミュニケーションにある、と考えます。インターネットを利用する際にも、科学技術の進歩がもたらす目先の新しさに心を奪われるだけではなく、情報をどれだけ自分の知識とすることができたかと常に自問自答したいものです。

補記: SIL 社の連絡先は下記の通り。

SIL Printing Arts Department
7500 West Camp Wisdom Road Dallas, TX
75236 USA
TEL +1 (214)709-2495 FAX +1 (214)709-3387
Fred West Internet: Fred_West@sil.org
インターネット上のメーリングリストについては以下の書籍を参考にした。

Eric Braun(1994), *The Internet Directory*,
Fawcett Columbine, New York, ISBN: 0-449-90898-4

学会カレンダー(Ver. 2.0, '95)

1995年4月10日～13日	KB&KS'95, Second International Conference on Building and Sharing of Very Large-Scale Knowledge Bases, University of Twente, The Netherlands Contact: KB&KS'95, c/o Knowledge-Based Systems Group, Dept. of Computer Science, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, Fax: +31 53 339605, E-mail: kbks95-org@cs.utwente.nl
1995年4月19日～22日	PACLING '95, Pacific Association for Computational Linguistics 2nd Conference. The University of Queensland, Brisbane, Queensland, Australia. Contact: Roland Sussex, Centre for Language Teaching and Research, The University of Queensland, Queensland 4072, Australia. Phone:+617 365 6896, Fax:+617 365 7077. E-mail:sussex@lingua.cltr.uq.oz.au
1995年5月29日～30日	Tutorial on Automatic Dictionary Making, Brussels, Belgium Contact: Mrs. V. Vienne, ILMH, CTB, 11 rue d'Arlon, B-1040 Brussels, Belgium
1995年5月29日～31日	NLULP5, Fifth International Workshop on Natural language Understanding and Logic Programming. Fundacao Calouste Gulbenkian, Lisbon, Portugal. Contact: Gabriel Pereira Lopes, Department of Computer Science, Faculty of Science and Technology, Universidade Nova de Lisboa, Quinta da Torre, 2825 Monte da Caparica, Portugal. Phone:+351-1-295 3220, Fax:+351-1-295 56 41. E-mail: gpl@fct.unl.pt
1995年5月31日～6月1日	Value Analysis of Specialized Dictionaries, Brussels, Belgium Contact: Mrs. V. Vienne, ILMH, CTB, 11 rue d'Arlon, B-1040 Brussels, Belgium
1995年6月11日～25日	Summer Seminar: Methods and Tools for Electronic Texts in the Humanities, Princeton University, USA Contact: CETH, 169 College Avenue, New Brunswick, NJ 08903, USA Tel: (904)932-1384, Fax: (908)932-1386, E-mail: ceth@zodiac.rutgers.edu, Please include the word "Summer Seminar" in the subject line of any e-mail enquiry
1995年7月9日～13日	18th International Conference on Research and Development in Information Retrieval. The Sheraton, Seattle, WA, USA Contact: E-mail to <sigir95@u.washington.edu> to be added to mailing list: today
1995年7月11日～15日	ACH/ALLC '95, Joint Annual International Conference of the Association for Computers and the Humanities (ACH), and the Association for Literary and Linguistic Computing (ALLC). University of California, Santa Barbara. Santa Barbara, California, U.S.A.

平成7年度情報知識学会研究報告会・総会開催のお知らせ

下記の通り、平成7年度情報知識学会研究報告会および総会を開催いたしますのでご参加をお願いいたします

記

日時：1995年5月27日（土）10:00-18:45

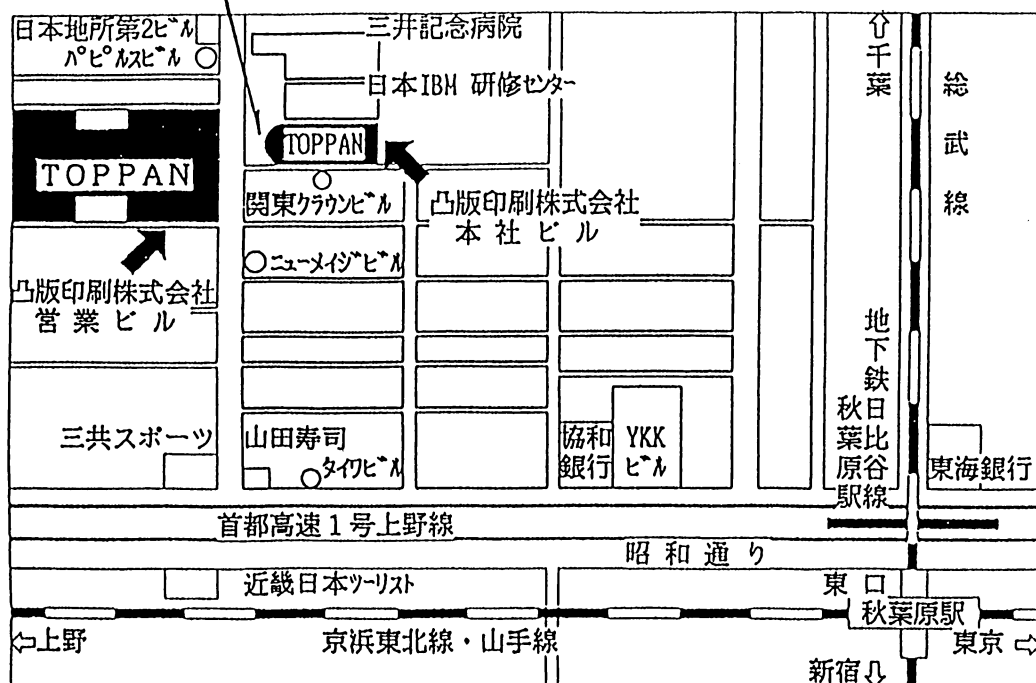
会場：凸版印刷本社1階ホール

プログラム

	A 会場	B 会場
10:00	〔概念と言語(I)〕 座長 芦崎達雄 (J I C S T) A1：専門用語研究における「意味」と「概念」 影浦峯 (学術情報センター) A2：概念構造構築のための概念モデル 平野朋成 (筑波大学) A3：単語の意味論的進化情報を付加した英語シソーラスデータベースシステムの構築 山崎達也 (筑波大学) ほか2名 A4：テキストからの類義語抽出手法とその評価 福島俊一、下村秀樹 (NEC)	〔専門分野知識〕 座長 菅原秀明 (理化学研究所) B1：オブジェクト指向設計法によるチベットの活字文字認識について 小島正美 (東北工業大学) ほか5名 B2：印刷された化学構造式の認識理解システム 伊藤尚樹, 中山堯 (神奈川大学) B3：非線形光学材料情報における意味関係の抽出と情報の構造化 宇陀則彦 (図書館情報大学) ほか3名 B4：蛋白質抗原sequential epitopeの特徴とPIR 志村純子 (理化学研究所)
12:00		
13:00	〔概念と言語(II)〕 座長 影浦峯 (学術情報センター) A5：用語集からの要素語推定の試み 小山照夫 (学術情報センター) A6：例文検索方式による多国語翻訳システムの実用化について 岩淵幸雄 (都築学園) ほか	〔マルチメディア環境〕 座長 中山堯 (神奈川大学) B5：マルチメディア型言語データベース(コーパス)の構築とその応用について 上村隆一 (福岡工業大学) B6：WDCMにおけるWorld Wide Webサーバの構築 宮崎智 (理化学研究所) ほか3名
14:00	〔社会知識〕 座長 小山照夫 (学術情報センター) A7：社会科学文献情報の生産水準の国際比較 岸田和明 (駿河台大学) 松井幸子 (図書館情報大学) A8：政府統計情報の公表形式の新展開 松田芳郎 (一橋大学)	〔文書構造・SGML〕 座長 田中洋一 (凸版印刷) B7：文書構造定義のためのSGML-DTD設計支援システム 小林尋文、石塚英弘 (図書館情報大学) B8：SGML文書の変換・再利用のための言語"AEsop" 高橋亨 (日立製作所)
15:00		
15:15	招待講演 座長 藤代一成 (お茶の水女子大学) 電子図書館の動向と課題 安達淳 (学術情報センター)	
16:15		
16:30	情報知識学会総会・懇親会	
18:45		

総 会 会 場

※ 本社ビル正面受付にお寄り下さい



編 集 後 記

「中世哲学研究の立場から、人文系学生のコンピュータ教育の現状から、そして経営の実践の報告、とさまざまな立場からお書きいただきました。またインターネットの体験記もお寄せ下さいました。そういえば、大震災でインターネットが、全ての電話回線が断たれた中で劇的に生き残っていたというニュースは記憶に新しいと思います。しかしそれとともに、我が国ではいまだそれが主たる情報網として、首相官邸を含む中央でも十分に活用される段階になかったことも分かりました。なにが、それを危機に際しての生命線ではなく、依然として来るべき未来のメディアに留めたのでしょうか。ともあれ、今や全ての人にとって、コンピュータはブラックボックスではなく、情報科学は現代の啓蒙ではもはやなく、光ネットがいわゆる共在 (Mitsein) の下支えとなるべき時代が到来していることだけははっきりしました。今回、初めて編集作業を受け持ち、慣れない為に執筆者の方々にも印刷担当の方にもご迷惑をおかけしたと思いますが、ご協力を心から感謝致します。」 (岡本)

5月27日(土)開催の情報知識学会研究報告会・総会のプログラムが決まりました。今回も充実した内容です。どうぞ奮ってご参加下さい。

2週間余りかけ北米の大学をいくつか訪問してきました。どこに行ってもインターネットが話題になりました。インフラ整備も大事ですが今後は情報サービスの中身が問題となりそうです。なお帰国が3月下旬になったため31号の発行が遅れました。深くお詫び申し上げます。

(長瀬)