

情報学研究連絡委員会報告

—情報学とその課題—

平成6年6月27日

日本学術会議

情報学研究連絡委員会

この報告は、第15期日本学術会議情報学研究連絡委員会の審議結果を取りまとめて発表するものである。

- | | | |
|-----|-------|-------------------------------|
| 委員長 | 尾関 雅則 | (日本学術会議第4部会員、(財)鉄道総合技術研究所理事長) |
| 幹事 | 田畑 孝一 | (図書館情報大学教授) |
| | 根岸 正光 | (学術情報センター教授) |
| | 藤原 譲 | (筑波大学電子情報工学系教授) |
| 委員 | 大塚 雍雄 | (新潟大学農学部教授) |
| | 開原 成允 | (東京大学医学部附属病院中央診療情報部長) |
| | 杉田 繁春 | (国立民族学博物館教授) |
| | 長尾 真 | (京都大学工学部教授) |
| | 早川 武夫 | (神戸大学名誉教授) |
| | 米田 幸夫 | (東海大学開発技術研究所教授) |

情報学とその課題 (要約)

背景および趣旨

人類の発明した言語, 文字, 紙, 活字などに加え, コンピュータと通信は, まさに来るべき, 情報化社会の基礎をなす技術である. 情報の管理, 処理及び伝送を取り扱う通信・情報工学が, いわば情報の通路と容れ物を対象とする学問であるのに対し, 中身である情報そのものの特性や構造を究明する情報学の進歩に対する期待は, 今後, 複雑化, 高度化の一途を辿るであろう情報化社会の正しい理解と方向づけのためにも極めて大きい. ここに, 情報学の確立と発展が望まれる基本的な意義が存在する.

本報告書では, 未だ十分に確立されるに至っていない情報学の内容と領域を考究し, 現時点で情報学の中心的な課題と考えられているものを示し, 更に将来の情報化社会において, 大きな貢献があると考えられる情報技術及び, それに携わるプロフェッショナルの在りかたなど, 新しい社会規範の必要性とその中心命題についても触れた.

1. 情報学概要

情報学とは, 「情報の本質に関する理論や知識を体系化すること, さらにそれらの応用として思考活動の質と効率の向上を図ること」つまり, 「情報とは何か」, 「情報をどう使うか」を追求することであるとする.

つまり情報を有効に活用するためには, 情報や知識がどういうものであるかを知る必要がある. それが情報の学問であるから, 情報学の目的は情報の特性と理論の体系化, 情報に関する技術, 手法の開発及びそれを具体的に各分野の情報, 思考活動への応用することになる.

情報学でまず問題になるのは, 情報, 知識, データなどの定義である. 情報とは最も広義では「認知とか思考の対象となる実体についての認識内容」であり, 普通の意味で言われる情報は全てこれに入る. 次に知識とは一般的には情報と同じに使われることもあるし, 情報処理, 特に人工知能の分野では一定の形式化された知識を指すので, 具体的にはプロダクションルールとか1階述語論理で表現されたもの, またはその延長上にあるものということになる. ここでは最広義と最狭義との中間になる, 「体系化された情報」という意味で使うことにする. 次に情報はいろいろな形で記述され, 表現されるがその「最小単位」をデータと

いう。またその「集合」もデータという。以下はこれらの定義に従うことにする。

情報の基本的特性を挙げると、a. 媒体依存性, b. 記述, 表現の多様性, c. 様相性 (Modality), d. 非加算性, e. 階層性 (入れ子構造), f. 相対性, 双対などがある。このような情報の特徴とそれに関連する課題についてこれまで多くの研究がなされているが、未解決の問題も山積して、学問としてはまだ緒についたばかりである。

情報学は全体としてどのような専門分野から構成されているかを自然科学の側から整理してみると、理論的、実験的な側面および応用的の3つの領域に分けられる。まとめると情報学は下のようになる。

情報学

1. 背景, 目的 : 領域-大量情報の特性、資源化、操作に関する理論と基本課題, 背景, 歴史, 概要, 基本概念の定義
2. 情報解析 : 特性解析-属性, 媒体, 動態
意味解析-物理関係, 概念関係, 論理関係, 従属関係など
3. 情報構造 : グラフ, ハイパーグラフ, 木構造-分類, 網構造
拡張ハイパーグラフ, 双対, 入れ子, 部分共有, 動的構造など
情報空間モデル化
4. 完全性制約 : 空値問題, 実在制約, 識別, 同定
5. 媒体 : 物理媒体, 論理媒体, 表現媒体, 記録媒体, 表示媒体, 通信媒体など媒体変換
6. 情報記述 : 属性空間, 記述項目, 差分記述, 様相性
情報表現 : 媒体依存性, 多様性, 多義性, lattice
情報表示 : 多元媒体
7. 情報構造化 : 意味関係構造化, 自己組織化, 学習
概念構造 : シソーラス, 構文解析辞書, 定義辞書
論理構造 : 述語論理, 様相論理, ファジー論理
物理構造 : アドレス, 索引, 所在, ファイル
8. 構造操作 : 記憶構造, 直接アクセス, 構造経由アクセス, 同型性, 準同型性, 拡張関係型操作-関係グラフ, 抽象化 (汎化, 集約化)
9. 意味処理 : 内容検索, 演繹推論
類似性処理 : 共有概念, 類似度-帰納, 類推, 仮説推論, 連想, 発想など
10. 応用システムおよび展望 : 自己組織型情報ベースシステム, 人工頭脳など

最近マルチメディアが注目されているが、ハイパーメディアはマルチメディアの有力な利用形態のひとつである。次世代のハイパーメディアに展開するために、解決すべき問題としてHalaszが87年と91年に改訂しCACMに発表した問題は大規模情報の管理と構造の適切な処理に要約できる。基本的には利用者向きで情報媒体の面からも望ましい大量の多様な情報の利用のためには情報の特性に即した新しいモデルに基づくシステムの開発が必要である。そのことによって意味関係の構造化を用いて新しい情報システムの展望が得られる。

2. 情報学研究の現状と展望

(1) 言語情報

情報を担っているものは言語だけでなく、図面、写真、映像、その他種々のものがある。それらは表現しようとする情報内容によって使い分けられる。情報はまたいかに客観的に相手に伝えられるかという立場からその媒体を考えることもできる。機械の設計図面などは世界中で共通的に理解されうるものである。それでは言語はどのような情報の表現に適しているのだろうか。あるいは情報学における言語の位置付けはどうであろうか。それは次のように考えられる。

(i) 言葉は思想を表現するための最適の媒体である。

(ii) 言葉は誰にでも理解でき、人による理解の相違を最も小さくし、的確な情報を伝達することのできる媒体である。

(iii) 歴史的に見て、人類の知的財産が言語によって最も多く表現され、また蓄積され今日に伝えられている。

(iv) 今日の情報技術においては、言語が最も安価に、最も容易に扱える媒体である。

このような理由から、言語技術は情報学の中で重要な位置を占めていることが分かる。

言語技術は字面処理、文解析、文生成、翻訳、テキスト情報の圧縮と分類などの段階があり、課題としては、テキストデータの蓄積、辞書、解析などのソフト等の基盤整備、言語理解のための知識辞書の作成から電子図書館の実現までがある。

(2) 情報標準化

情報に関する標準化には、科学技術会議が総理大臣に答申した「科学技術情報流通に関する基本政策」に従った情報流通技術の標準化(SIST)と、従来展

開されてきた日本工業規格（J I S）での情報処理技術の標準化の二つの流れがある。いずれもそれぞれの自的に沿って、情報の流通・処理における技術の整合性を高めることに主眼を置いている。この経緯には、科学技術振興における支援活動として早くから進められてきた前者と、機器の進展に比して産業化の遅れた情報活動に対する後者の認識の差が認められる。経緯はともあれ、両者とも国際標準への整合を掲げており、その点で共通の基盤をもっているといえることができる。強いていえば、後者が工業技術の延長上で「情報」を眺めているのに対し、前者は科学技術が内包する「情報」に視点を置いているところに相違がある。また、前者が20年の実績を持っているのに対し、後者は国際標準化機構への国家対応の公式規格であることに、それぞれの特徴を有している。制度的には、前者が科学技術庁科学技術振興局の発行文書であるのに対し、後者は工業標準化法に基づく法律の認知を受けている。内容的には前者が書誌情報および情報生産を対象としているのに対し、後者は用語・略語・記号・符号などが法律にその対象として例示されている。両者の境界については微妙なところがないでもないが、S I S Tには抄録作成・参照文献記述・レコードフォーマット形式など書誌情報に関するものと、学術雑誌構成・学術論文構成・科学技術レポート様式・予稿集様式などの情報生産に関するものの2種となる。I S O / T C 4 6（情報ドキュメンテーション）の制定した国際規格のなかで、J I Sとして国名コード、I S B N、I S S N、ドキュメンテーション用語などが制定された。

（3）全文データベース

いわゆるデータベース・サービスは、文献の書誌的データと要旨を収録した二次文献データベースを中心に発達してきた。これには、従来の抄録誌の編集作業が電算化され、そこで得られる電子化ファイルが、当初は副産物的位置付けにあったが、次第に主製品としてのデータベースに転化してきたという背景がある。一方、数値情報系のデータベースは、株価の即時配信システムという、データベース的ではないオンラインシステムからはじまったが、その後、データの蓄積機能を取り入れてデータベースを構築するようになった。また統計情報関係では、統計データのデータベース的整備と、これに対する分析プログラムの組合せでサービスするシステムが早期に事業化されている。全文データベースは、上述の二次文献データベースの進化の延長上にあるものである。

全文データベースの発達は、画像・音声マルチメディア型データに関する情報技術のみならず文字検索にしてもシグナチャーファイルなど全文対応の技術を必要としている。もっとも、高速・大容量・低価格のインターネットの普及に伴っ

て、オンライン系でも、画像・音声を含むハイパーテキストの普及の兆しを見せている。また、既存出版物の各頁を画像として蓄積・配信することも、インターネットの普及によって現実的になっており、このような全文データベースの集積と配信を総合したシステムを「電子図書館」と称して、その実用化にむけた開発計画が米国ではいくつも試みられている。

(4) マルチメディア

我々の日常はマルチメディアの世界である。文字、画像、音響など様々な情報メディアに囲まれて生活している。ところが1980年代に入ってマルチメディアという言葉が新聞や雑誌広告などにしばしば見られるようになった。これはコンピュータや通信の世界において単に文字や数値だけではなく、静止画のみならず、動画や音響も扱える様になりつつあることを反映して特に強調されてきた表現である。

画像や音響データはそれをデジタル化すると膨大なデータ量になる。一枚の写真も文庫本10冊分くらいの量になる。それを数千枚数万枚を扱うとなると大変な量である。従来のコンピュータではシステムが大きくなり過ぎた。ましてやパーソナルコンピュータでは不可能であった。ところが大量のデータを蓄積することが出来る媒体、光ディスクの出現によって事情が一変した。数百メガバイトのデータを蓄積することができる光磁気ディスクやCD-ROMをコンピュータに接続することによって従来の計算機のメディアの世界が飛躍的に広がって、我々の日常的なメディアの世界に近づいてきたのである。

入力用スキャナー、データ圧縮技術、光通信、衛星通信などの技術の進歩に支えられて応用が広がりつつあるが、データ間の関係付け、言語による記述法などの課題も多い。

(5) 情報自己組織化

高度な思考機能に対応する情報処理をするには情報の意味処理をする必要がある。ところが意味の関わる問題の多くは未解決である。例えばデータベースや知識ベースでは個別実体(Distinct Entities)の集合として、外延的にデータや知識を対象としていることである。また類似性というのも重なりがある概念の関係であるから同様である。あるべき情報が欠落している空値問題はさらに意味処理が困難である。これらが情報の管理や、知識の獲得の困難さの原因となっている。表現の多様性も情報の記述、目的、内容に応じて大きく変化する。分類や表現の多様性は情報の表現の本質的な性質であるので意味の処理が困難になるのである。

この様な問題の解決策として脳における機能と同じように情報の意味的關係を自動的に構造化する自己組織化の研究が注目されている。

脳における学習に対応して、概念や情報間の意味的關係を抽出して情報の組織化を行う。概念に内在する關係は概念構造としてシソーラスを構築する。論理關係は、原因—結果、理由—結果、要因—結果を主としてタキソノミーの形で組織化する。原情報の所在、書誌情報、アクセス情報などは物理構造として構築する。メディア変換、意味關係抽出等によりマルチメディア型原情報の概念構造、論理構造、物理構造などの自己組織化を行ない演繹推論、帰納推論、類推などの可能な人工頭脳型システムを設計する。研究に必要な情報の動的構造の記述操作のためには新しい型の情報構造を持つモデルSSR(Structured Semantic Relationship)の開発がなされている。そのためデータベース、知識ベースおよびハイパーメディアなどの要素技術を大幅に拡張し、新しいモデルで統合的に記述表現、操作し、思考支援できる機能を持つ情報ベースシステムの設計と、基礎となる理論の研究が行われている。

(6) 国際全文データベースサービス

学術文献情報に関するデータベースサービスは長くこれまではいわゆる二次情報を提供するサービスであった。利用者が指定したキーワードを基にデータベースを検索し、論文名やその書誌的事項、あるいはその抄録を提供するというものである。データベースの構築者は、学術文献のたゆまぬ収集と蓄積に努め、それに基づく利用者へのサービス提供をビジネスとして維持してきている。

最近になって、学術文献の一次情報、すなわち論文の本文そのものを提供する全文データベースシステム構築の国際的な実験プロジェクトが欧米を中心に開始されている。このシステムの特徴は次のようである。(1)論文・記事の本文を提供、(2)国際学術ネットワークを介し、多くのサイトが自由に参入して相互にデータベースを提供、(3)全文検索、ハイパーテキストなどの新検索手法の導入、(4)記述の国際標準化、(5)文字情報のみならず画像、音響情報を提供、(6)国際学術ネットワークを介し、どの国の誰でも利用可能。

これらのプロジェクトの例として、WAISとWWWの二つがある。これらはいずれも国際学術ネットワークであるインターネットの上に構築されていて、インターネットに接続したワークステーションから誰でも利用できる。

3. 専門領域における情報学の現状と展望

専門領域として、ここでは医療情報、判例データベース、農業情報、博物館情報を取り上げる。

日本の医療情報の分野は、実用面では一部で高い水準の保健医療情報システムが稼働しているが、全体的なポリシーが欠けているために、システムが相互に連携しない形で作られ、多くの無駄や非効率な点が見られる。また、データに基づき客観的な意思決定をする習慣がないため、情報システムが医療機関の運用効率の追及に使われ、情報を収集・分析して意思決定に使うという認識が少ない。

一方、研究面では、日本で生まれたイノベティブな医療情報関連技術は少ない。医療関連の新しい技術は、ソフトウェアの形で生まれることが多いが、日本ではソフトウェアの価値に対する認識が少ないことも上記の理由の一つであろう。

今後重要な政策的課題として、ここでは次の7つをあげる。(1)医療情報政策を討議する場の創設、(2)ソフトウェアの価値の認識、(3)医療情報関連技術の標準化の促進、(4)情報を利用することに対する経済基盤の確立、(5)情報技術応用の障害となる法令の見直し、(6)一般人への医療情報の提供の原則の確立、(7)情報技術を用いた国際医療協力。

判例検索は乾草の山の中から1本の針を探すのにも似た大変な作業である。判例法を主要法源とする米国においては、年々連邦や諸州その他の法域の最高裁以下の諸裁判所から流出されるおびただしい数の判例がデータベース化されている。判例データベースは、当初は判例への手がかりを集大成したものであったが、現在では判例の全文が入力されている。検索方法は、全文を対象としたKWIC

(Key Words in Combination)方式が採用されている。key wordsは全文の中の前置詞、接続詞、冠詞、人称代名詞などを除いたすべての語である。

農業情報の特徴について考察する。農業研究の分野は、生物の改良・機能の解明、生産の場としての物理的・生物的環境の解明・制御、環境に人為的に投入される資材の開発・改良・管理、農業基盤整備、機械化、施設化・農作業の効率化、農業・農村計画、農家経営の改善、農業生産動向の解析・予測、食品の安全性・高品質化など多様であり、研究手法も自然科学から人文科学まで広範囲でしかも扱われる情報も多種・多様である。さらに、最近では農業農村の活性化、環境保全、国土保全機能等農業の持つ多面的機能の向上が重要な研究課題となっている。

農業情報の特徴と関連して農業分野における情報研究の課題をいくつか挙げる。(a)巨大情報の管理・処理手法、(b)システム科学的手法、(c)データベースの開発、(d)農業農村の高度技術化、高度情報化。

現在、農業における情報研究は情報の計測・収集手法、管理・蓄積手法、情報システム化手法など広範囲に実施されているが、必ずしもうまく進んでいるとは言えない。それは情報学専門の研究者がいなかったこと、組織的な研究体制がなかったこと及び情報研究の成果を正当に評価する認識に欠けていたことによるものと考えられる。

博物館情報について考察する。普通博物館という名称は考古資料や生活資料を陳列している場所として狭い意味に使われている、しかし美術館、水族館、動物園、植物園、科学館、なども博物館の範疇である。そこでは対象の違いはあっても、さまざまなモノを集め、それをある種の体系に従って展示をし、そのモノの世界をより良く理解させるようにする施設である。

普通どこの博物館でもその館が対象にしているモノに対して情報カード、あるいは資料カードが用意されている。そこにはいろいろな項目が書かれている。名称などの基本情報はカードに記入されているが、そのほかの情報となるとほとんど記入されていないのが現状である。ましてやコンピュータ化がなされている場合は少ない。情報カードの作成ができていない原因のひとつは、どのように分類するかで行き詰まってしまうからである。分類体系の確立は困難であるから、コンピュータを利用した別の検索体系を模索する必要がある。例えば、情報カードに含まれている文字列すべてが検索の対象になるから、断片的なデータであっても活用可能なのである。

大小さまざまな博物館で情報化が進んでいない要因として標準化ということにこだわっているところがある。図書の分類体系と同様の標準的な項目をすべての博物館で共通に利用できるのを待っているというケースがある。標準化はそれができるに越したことはない。なかなか足並みが揃うものではない。それを待っていたのでは資料がどんどんたまる一方である。各館独自のデータフォーマットでも検索できる体制にあれば、他の博物館とのネットワークによって相互利用ができるのである。記述データだけではなく博物館では立体的な画像を含めた画像データが重要になる。

4. 情報化社会の規範

情報化社会の規範について、情報経済学、知的財産権、情報倫理綱領の観点から論ずる。

経済活動における情報の役割の重要性が増大しつつある。そこでは情報として価格情報のみならず、財の品質や特性に関する情報なども扱われる。一方、情報を

経済財の一種として、その特殊性を分析し、情報財の価格、需要と供給、市場、情報産業などについての解明を目指している。

経済財としての情報は、その特質として、排除不能性（情報は安価に複製でき、こうした利用を不正として排除するのが困難である）、協同消費性（もう一人が情報を得るについて、追加的費用を要しない。情報は見ても減らない）がある。そこで、こうした情報財取引の不正化や、情報産業の振興を図るべき経済政策などについて、例えば、複製防止装置や複写料徴収制度の経済効果を検討し、また産業関連表を情報活動に着目して組み替えるといった試みがなされている。

知的財産権制度は、情報化社会において新しい対応を迫られつつある。著作権制度に焦点を当てて問題点を列挙する。(1)オリジナルとコピーの区別の無意味化、(2)複製概念の変質、(3)人格権の形骸化、(4)オリジナルの価値の低下、(5)公的情報の権利化、(6)私的使用の増大、(7)情報の自由流通、(8)権利の集中管理システムの立ち遅れ、(9)著作物の国境越境流通。

およそ職業人には高度な専門知識・技能の修得と、これを用うるに当たっての高い行動規範の遵守とが要求される。この点で専ら利を追求する手職人や商人と区別される。専門知識・技能は高等専門教育や資格試験の問題であり、行動規範は倫理綱領の問題である。

コンピュータ時代に入って登場した新しい型の情報関係者のうち、情報技術者は上記の両資格要件を満たすべき職業人と考えられる。かれらのための倫理綱領の作成は現下の急務である。既成の諸職業倫理綱領の具体例を、米国法曹協会における法曹倫理綱領を例にとってその問題点を検討した。

(1)綱領は理想や目標の宣言にとどまらず、具体的で実践的な行為規範を盛り、(2)それらの行為規範は制裁を伴い、一部は法規規範と重複する。(3)具体的で詳細だから、情報倫理綱領は部門ごとに異なることになる。

法は服従の意志の有無を問わず実力で強制するが、倫理は元来良心の問題として自発的服従を予定するから、実効の確保には世人の批判・非難による以外ない。職業団体がある場合、違反者に注意・警告の上除名する。その団体に権威があれば、除名は本人の社会的な信用の失墜として制裁になる。

I F I P (International Federation for Information Processing) においても情報倫理綱領をめぐる課題の検討を始めている。国際機関が全加盟学・協会に受け入れられる規約を作成することは困難であるが、加盟各国の学・協会がそれぞれの国情にあったものを作成する際、守るべき原則を示すことは可能であると考えている。

5. わが国における情報学の課題

わが国における情報学の課題として、ここでは情報資源の整備、日本語情報の収集と提供、情報流通の整備、著作権の権利処理、情報資料の蓄積交換機構の観点から取り上げる。

情報技術の高度化により、「情報」を体系的に収集し、加工し、処理し易いように整備・蓄積することができるようになり、それらを利用価値の高い「資源」と呼ぶことがふさわしい時代となった。しかしながら、わが国における情報資源化は欧米に比べ遅れており、国内で利用できる資源のうち、国産のものはたかだか4分の1にしか過ぎない。

情報資源の整備には長年にわたる多くの知的労働と資金が必要である。そのために次の方策が考えられる。(a)研究者や研究機関など情報発信者や防災、環境、医療機関など各種情報収集機関への情報資源化のための財政的、人的、技術的な支援の強化、(b)大学や国公立試験研究機関などが積極的に研究情報を公開できる環境の整備、(c)情報資源化のための情報表現形式の標準化の推進、(d)わが国の主体性の確保と国際的役割を果たすための基盤となる情報を整備する公共的な情報資源化専門機関に対する助成の充実。

日本語情報への需要の源は、日本語教育、言語学教育、コンピュータによる日本語処理の三つである。従来は研究面からの需要という、ごく限られたものであったが、国際化の広がりの中から外国人からの実用日本語への要求、また日本語ワープロの普及、機械翻訳を代表とする日本語処理ソフトウェアの作成などにより、最近、需要が高まってきている。

このような需要に対応するため、供給側として大規模な電子化辞書の編纂などの形でいくつかの官民の機関の努力が払われているが、まだまだ不十分である。日本の国家情報基盤そのものである日本語情報の収集整備に、より真剣にとり組む必要がある。

流通すべき情報として、灰色文献 (Grey Literature) と呼ばれるカテゴリの文献がクローズアップされつつある。灰色文献とは通常の出版・流通の経路で扱われていなく、またそれについての検索手段が整備されていないので、入手が困難な文献を指す。これにはテクニカルレポート、会議資料、学位論文、大学出版物、官公庁資料などが含まれる。正規の流通経路にある学術雑誌は研究成果が得られてから出版までに時間がかかり、現場の研究者にとってはテクニカルレポート、会議資料などがより重要となる。また官公庁資料は民間の企業活動にとって

貴重であるがなかなか入手できない。灰色文献の収集と蓄積、検索、配布については、これまではその二次情報をオンライン情報検索システムやCD-ROMで提供され、ものによってはその一次資料のコピーを郵送あるいはFAXで配布されている。しかしながら、対象とする文献は限られたもので、またその事業の採算性に難点があり、このままでの体制では問題の解決にほど遠い。最近になって、問題の解決を国際学術情報ネットワークを介した全文データベース（一次資料）の相互共有、所在情報サービス、一般利用者への文献本文（一次資料）の提供に求める動きが活発になってきている。

すなわち、国際学術情報ネットワークに、灰色文献の生産者となる組織が加入する。各組織は自身が生産する文献を格納する全文データベースを維持し、ネットワークからのアクセスに応じる。望むらくは、各国の産官学のすべてがそのような組織として加入する。このネットワークにはもちろん一般の利用者がアクセスでき、文献本文が電子的に入手できる。

この方式を進めるにあたって解決すべき点がいくつかある。(1)文献データの国際標準化、(2)マルチメディア情報、(3)所在情報、検索手法の開発と整備、(4)著作権使用料、(5)一般利用者からの料金徴収方法、(6)データベース提供者への利益還元方法、(7)ネットワーク運営・財政の管理。

情報の電子化に伴って、情報の処理・提供方法は飛躍的に多様化し、この傾向は、昨今多用されるマルチメディアという用語に象徴されるように、現在ますます強まっている。そして、この種の応用を研究的にまた事業的に推進しようとする立場から、著作権の処理が重大な制約要因とみなされるようになってきている。こうした事態は、これまで著作権の問題に直接関係しなかったような企業、研究者の参入、すなわち、彼らにおける著作権処理に関する無知、誤解に由来する部分が多いようである。一方、著作権者の側においても、マルチメディアなどの新しい公刊形態に対する誤解・不安という要因がある。これらは当局の啓発活動によって逐次解消されるべきであり、各般の措置がすでに講じられているところであろう。しかし、啓発のみによっては解決し難い部分もあり、法的あるいは制度的な方策も検討されつつある。

情報学の研究においては、研究の素材あるいは実験試料として、各種の著作物を多種・大量に利用することが不可欠である。現状では、何分新しい事態であるため、研究者側と著作権者側の双方に上記のような無知や誤解があり、個別的な利用交渉は一般にはかばかしくないようである。情報を保護する制度によって、情報学の進展が阻害されるとすれば、これは重大な矛盾であろう。

元来、著作権法には、例えば図書の売買契約に際して、これに並行して、その

内容の利用方法を個別に著作権者と契約交渉するという煩瑣な過程を省略するための、著作権利用約款の役割があると考えられる。このような著作権処理の定型化、簡素化によって、従来型の著作物の流通は良好に維持されている。この点に鑑みると、新しい情報技術に対応した新しい形態の著作物に対しても、その利用契約に関する手順の明確化、標準化、簡素化が望まれる。その初段階として、著作権の所在をデータベース化した「著作権権利情報集中機構」の設置を考慮することも考えられる。

著作物は著作権者の私的財産であると同時に、社会的な知的資産でもある。したがって、これが新しい情報技術により一層有効に活用されることは、著作権者にとつても社会全体にとつても望ましいところであり、著作権保護の意義もここにあると考えられる。この際、情報学研究を助長するような著作権処理制度、体制の整備が望まれる。

上述の著作権処理機構を、情報学研究振興の立場からさらに発展させると、これは、情報研究者間で共通に利用されるような情報資料を集中したセンターの設置という構想にいたるであろう。この種の機構では、研究者相互間での情報資料の交換を簡便に行いうるような機能がまず想定される。しかし、より重要であるのは、研究者社会の外部、具体的には企業などから多様な情報資料の寄託を受け、これを研究者間の共用に付すという機能であろう。研究者と企業等著作権者間の研究的データ利用契約交渉は難航することが多い。大学研究者は、もとより営利的利用を目的とするわけではないが、権利者側では前例のないこととて理解に至らず、商業的利用のために設定されるような高額の使用料を提示することもある。このような価格は研究者からすればほとんど禁止的なものであり、結局利用を断念するというような例が聞かれる。情報資料の公的な蓄積交換機構の整備は、このような事態を打開するために有効に機能すると期待される。

〔付記〕

本報告の作成に当たっては、下記の方々の御協力を得た。ここに感謝する次第である。

- 仲本秀四郎 (IRIS情報学研究所長)
- 名和小太郎 (新潟大学法学部法学科教授)
- 横井 俊夫 (日本電子化辞書研究所長)
- 神田 利彦 (日本科学技術情報センター技術開発室長)
- 黒川 恒雄 (工学院大学常務理事会参与)

—情報学とその課題—
(本文)

目 次

はじめに	尾関雅則
1. 情報学の概要	藤原 譲
2. 情報学研究の現状と展望	
2. 1 言語情報	長尾 真
2. 2 情報標準化	仲本秀四朗
2. 3 全文データベース	根岸正光
2. 4 マルチメディア	杉田繁治
2. 5 情報の自己組織化	藤原 譲
2. 6 国際全文データベースサービス	田畑孝一
3. 専門領域における情報学の現状と展望	
3. 1 医療情報	開原成允
3. 2 判例データベース	早川武夫
3. 3 農業情報	大塚擁雄
3. 4 博物館情報	杉田繁治
4. 情報化社会の規範	
4. 1 情報経済	根岸正光
4. 2 知的財産権	名和小太郎
4. 3 情報倫理綱領	早川武夫
4. 4 情報倫理綱領をめぐる I F I P の活動	黒川恒雄
5. わが国における情報学の課題	
5. 1 情報資源の整備	神田利彦
5. 2 日本語情報の収集と提供	横井俊夫
5. 3 情報流通の整備	田畑孝一
5. 4 著作権の権利処理, 標準化および流通	根岸正光

はじめに

人間は太古より言語、文字を所有し、更に中世に至って紙と活字を発明した。これらによって、人間は時空の隔たりを越えて、相互に意志や感情を伝達することを可能にし、原始社会は次第に文明社会へと変貌を遂げてきた。現代に及んで物質とエネルギー利用の進歩と共に、情報の伝達、保存、変換などの手段は、通信とコンピュータ技術の出現によって、画期的な発展を遂げつつある。

殊に、現在まさに進行中であるパーソナルコンピューティングと、パソコン相互の自由な通信は、物質とエネルギー技術を中心に発達してきた産業社会を、次の脱産業社会へと変化させる原動力であると考えられるにいたった。

コンピュータと通信は、まさに来るべき、情報化社会の基礎をなす技術である。情報の保存、変換及び伝送を取り扱う通信・情報工学が言わば情報の入れ物を対象とする学問であるのに対し、中身である情報そのものの性格や構成を究明する情報学の進歩に対する期待は、今後、複雑化の一途をたどるであろう情報化社会の正しい理解と方向づけのためにも極めて大きいと言わなければならない。ここに、情報学の確立と発展が望まれる基本的な意義が存在すると言わなければならない。

本報告書では、未だ十分に確立されるに至っていない情報学の内容と領域を考究し、少なくとも、現在のところ情報学の中心的な課題と考えられているものの内容を示し、更に将来の情報化社会において、大きな力を持つようになると考えられる情報技術及び、それに携わるプロフェッショナルの在りかたなど、新しい社会規範の必要性とその中心命題についても論究した。

この報告書が情報学の内容と、その社会とのかかわりについて正しい理解のために少しでも役立つことを願って、前書きとする次第である。

1. 情報学概要

(1) 序

情報学とは、「情報の本質に関する理論や知識を体系化すること、さらにそれらの応用として思考活動の質と効率の向上を図ること」つまり、「情報とは何か」、「情報をどう使うか」を追求することであるとする。

先ず人間の脳における活動について考える。優れた思考機能を持つ頭脳の表現として博識という言葉がある。これは物理的に言えば情報の量が多いということに相当する。また頭脳明晰であるとか頭の回転が速いということも重要な機能であり、これは応答の速度が早く、しかもその結果の精度が高いということになる。これらは一応定量化ができる機能であるが、これだけでは高度な知的活動に対応しているという感じがしない。発想、創造などというレベルになると本当の意味で価値を生ずる機能といえる。

それにはどのような処理が対応しているかという、情報を収集、解析し、検索、数値計算、演繹推論のみならず、帰納推論、類推、仮説推論、連想や評価をすること、さらにその延長上に問題解決、意志決定、創造、評価などがある。

歴史的にみて「思考支援の方式」としてこれまでどのような手段を持っていたのかということを見ると、簡単なそろばんや計算の手助けになる道具は人間の歴史と同じくらい長い歴史があり、基本は加算であるけれども、四則演算ができる道具として各種の形のものがある。そのことを計算機ではスイッチングの機能で行わせる。四則演算が基本であることは同じであるが、高速大量の処理ができることで、論理演算も可能となっている。最近コネクションマシン、ニューロ計算機といったようなもので分類や学習が有る程度できるようになった。

さらに高度な思考機能的処理になると、意味の処理、内容を把握する必要がある。単純な数値や符号の処理だけでは済まなくなる。もしこういうことができれば機能として人工頭脳ができることになる。ところが意味の理解や処理は、現在までの理論や技術では極めて限定された範囲を除けば未解決の問題であり、とくに情報の記述、表現を含め情報の諸性質を把握することが必要とされている。

つまり情報を有効に活用するためには、情報や知識がどういうものであるかを知る必要がある。それが情報の学問であるから、情報学の目的は情報の特性と理論の体系化、情報に関する技術、手法の開発及びそれを具体的に各分野の情報、思考活動への応用することになる。

情報学でまず問題になるのは、情報、知識、データなどの定義である。情報と

は最も広義では「認知とか思考の対象となる実体についての認識内容」であり、普通の意味で言われる情報は全てこれに入る。次に知識とは一般的には情報と同じに使われることもあるし、情報処理、特に人工知能の分野では一定の形式化された知識を指すので、具体的にはプロダクションルールとか1階述語論理で表現されたもの、またはその延長上にあるものということになる。ここでは最広義と最狭義との中間になる、「体系化された情報」という意味で使うことにする。次に情報はいろいろな形で記述され、表現されるがその「最小単位」をデータという。またその「集合」もデータという。以下はこれらの定義に従うことにする。

情報が持っている特徴と、それらに関連する課題の例を次に示す。情報は数えられるものとして扱われることが多いが、本来可算集合ではないということが第一の特徴であり、識別子の設定や管理上の問題となる。次に計算機では2値論理が処理し易いが、計算機から見ても論理的に見ても厄介な多値論理が情報の本質である。つまり多数の同意語のあるのが用語の基本的な特徴である。また表現されたものに多義性があるが曖昧性を生ずる。例えば英和辞書を引くと、一つの英語に対して日本語が一つだけ対応しているという言葉は殆ど無い。通常非常に沢山の種類の訳語が書いてある。さらに意味の表現の他に表現の意味解釈の問題があり、それは計算機では情報そのものを扱っているのではなく、媒体上に記述表現されているものを対象としているということから生じている。これを情報の媒体依存性という。

このような情報の特徴と課題についてこれまで多くの研究がなされているが、未解決の問題も山積していて、学問としてはまだ緒についたばかりである。

(2) 情報学の専門領域

情報学は全体としてどのような専門分野から構成されているかを自然科学の側から整理してみると、理論的、実験的な側面および応用的の3つの領域に分けられる。

(2. 1) 理論情報学の構成

理論的な分野は上で述べたことから、まず情報の解析特に構造の解析方式であり、次に情報の表現や構造の意味対応などのために必要な理論的骨核としてモデルがある。また歴史のある分類の可能性と分類の手法、それから媒体に関連して記述表現の多様性、その取扱い方、情報の時間的、空間的、意味的变化、管理の可能性、限界、情報操作の手法などが情報学の基本対象である。

(2. 2) 実験情報学の構成

実験的な分野としては、実際の情報を対象として情報の特性、情報の量、情報の質、情報のキャラクターゼーション、情報の資源化、管理および典型的な情報媒体の要素であるターミノロジやシソーラス、辞書、日本語、マルチメディア、それらの構築、特性、操作処理などが実験的な情報学の領域である。

(2. 3) 応用情報学の構成

理論や実験が進めば応用も具体的に展開できるわけであり、典型的な情報検索の手法は確立されており、数多くのシステムが開発、提供されている。もう少し情報を高度に加工して付加価値を図ること、情報の伝達法として従来からの印刷物とオンラインデータベース、バッチ型データベースや知識ベースなどの位置づけと展開、さらに学習、類推、発想などを実現し、最終の目的である問題解決、意志決定、評価、人工頭脳までも一応応用情報学の対象に含まれる。

以上まとめると情報学は下のようになる。

情報学

1. 序 : 領域—大量情報の特性、資源化、操作に関する理論と基本課題、背景、歴史、概要、基本概念の定義
2. 情報解析 : 特性解析—属性、媒体、動態
意味解析—物理関係、概念関係、論理関係、従属関係など
3. 情報構造 : グラフ、ハイパーグラフ、木構造—分類、網構造
拡張ハイパーグラフ、双対、入れ子、部分共有、動的構造など
情報空間モデル化
4. 完全性制約 : 空値問題、実在制約、識別、同定
5. 媒体 : 物理媒体、論理媒体、表現媒体、記録媒体、表示媒体、通信媒体など媒体変換
6. 情報記述 : 属性空間、記述項目、差分記述、様相性
情報表現 : 媒体依存性、多様性、多義性、lattice
情報表示 : 多元媒体
7. 情報構造化 : 意味関係構造化、自己組織化、学習
概念構造 : シソーラス、構文解析辞書、定義辞書
論理構造 : 述語論理、様相論理、ファジー論理
物理構造 : アドレス、索引、所在、ファイル