

データベース構築促進及び技術開発に関する報告書

記事データベースアクセス用パイロットシステム構築

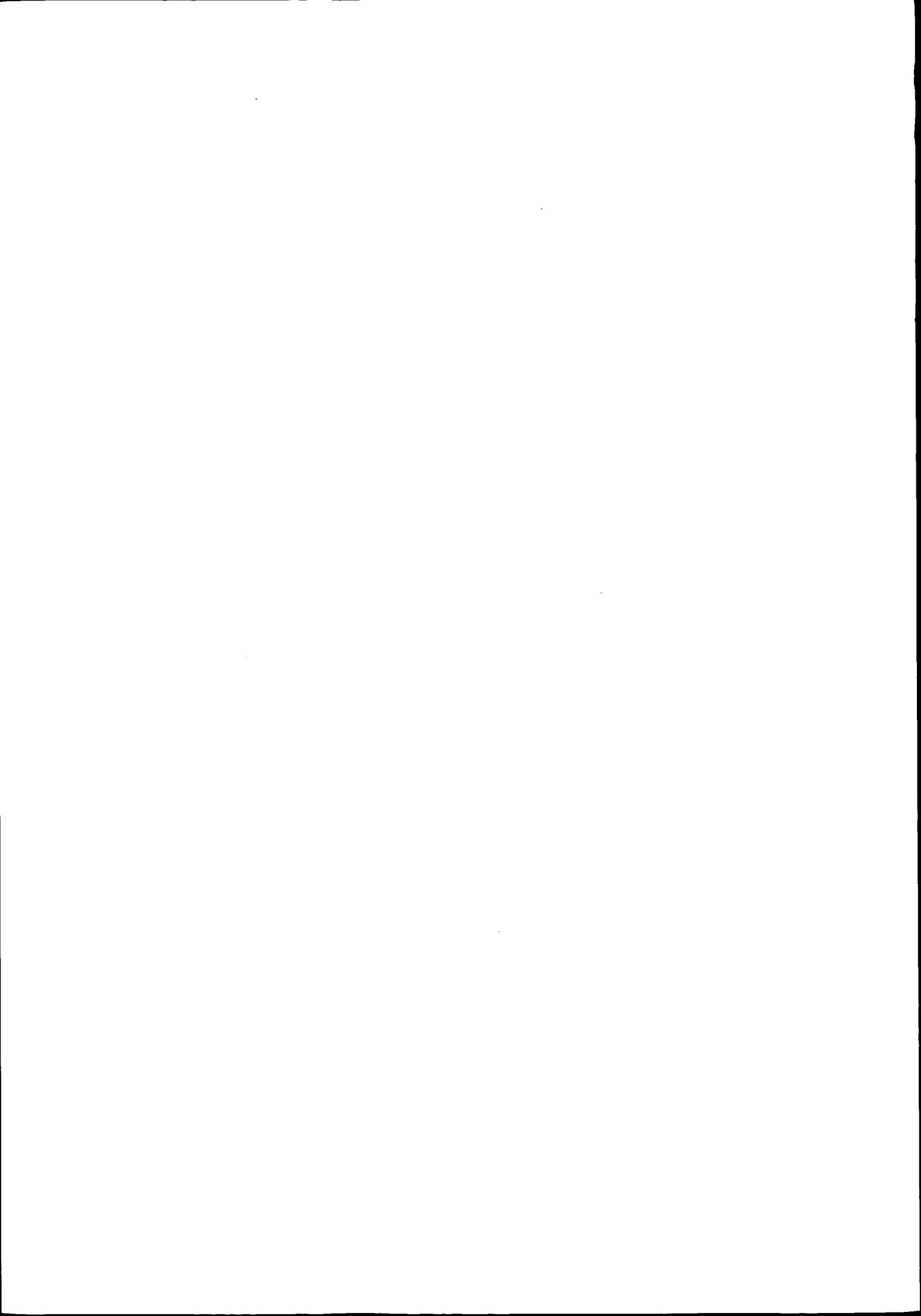
平成4年3月

財団法人 データベース振興センター

委託先 株式会社河北新報社

本報告書は、日本自転車振興会から競輪収益の一部である機械工業振興資金の補助を受けて作成したものである。





序

データベースは、わが国の情報化の進展上重要な役割を果たすものと期待されている。今後、データベースの普及により、わが国において健全な高度情報化社会の形成が期待される。さらに、海外に対して提供可能なデータベースの整備は、国際的な情報化への貢献および自由な情報流通の確保の観点からも必要である。しかしながら、現在わが国で流通しているデータベースの中でわが国独自のものは1/3にすぎないのが現状であり、わが国データベースサービスひいてはバランスある情報産業の健全な発展を図るためには、わが国独自のデータベースの構築およびデータベース関連技術の研究開発を強力に促進し、データベースの拡充を図る必要がある。

このような要請に応えるため、(財)データベース振興センターでは日本自転車振興会から機械工業振興資金の交付を受けて、データベースの構築および技術開発について民間企業、団体等に対して委託事業を実施している。委託事業の内容は、社会的、経済的、国際的に重要で、また地域および産業の発展の促進に寄与すると考えられているデータベースの構築とデータベース作成の効率化、流通の促進、利用の円滑化・容易化などに関係したソフトウェア技術・ハードウェア技術である。

本事業の推進に当たって、当財団に学識経験者の方々に構成されるデータベース構築・技術開発促進委員会（委員長 山梨学院大学教授 蓼沼良一氏）を設置している。

この「記事データベースアクセス用パイロットシステム構築」は平成3年度のデータベースの構築促進および技術開発促進事業として、当財団が株式会社河北新報社に対して委託実施した課題の一つである。この成果が、データベースに興味をお持ちの方々や諸分野の皆様方のお役に立てば幸いである。

なお、平成3年度データベースの構築促進および技術開発促進事業で実施した課題は次表のとおりである。

平成4年3月

財団法人 データベース振興センター

平成3年度 データベース構築・技術開発促進委託課題一覧

| 分野 | 課題名 | 委託先 |
|-----------------|---|---------------------------|
| 社 会 | 1 形態学的コメントを含む病理データベースのプロトタイプ作成 | (株)エス・ピー・オー |
| | 2 交通事故調査データベースのプロトタイプの作成 | (財)日本自動車研究所 |
| | 3 シルバーエイジの医療と福祉情報の実的な活用を目的としたデータベース構築 | 美崎高齢者福祉互助会 美崎生活館 |
| | 4 気象情報データベースの構築 | (株)エムテーエス雪氷研究所 |
| | 5 地下水情報データベースシステムの構築のための調査研究 | (株)パスコ |
| | 6 ファジィに関する文献データベースシステムの開発 | (財)日本情報処理開発協会 |
| | 7 大学におけるデータベース利用教育システムに関する調査研究 | 日外アソシエーツ(株) |
| | 8 マルチメディア型社会科用データベースの開発 | (株)新学社 |
| 中小企業振興 地域活性化 | 9 異分野研究のための知的オリエンテーション・データベースシステムの構築可能性調査 | (株)けいはんな |
| | 10 瀬戸内圏公共図書館の郷土資料データベースの構築 | (株)中国新聞社 |
| | 11 記事データベースアクセス用パイロットシステム構築 | (株)河北新報社 |
| | 12 商業調整支援データベースの構築 | (株)日本統計センター |
| | 13 地域流通最適化に必要なデータベースの構築に関する研究 | (財)日本ボランティア・チェーン協会 |
| | 14 情報源検索データベースのプロトタイプ作成 | セントラル開発(株) 情報図書館 RUKIT |
| 海 外 | 15 有価証券報告書の MT データ変換ソフト開発と英訳辞書作成 | コムラインインターナショナル(株) |
| | 16 海外の主要国際・国家規格データベースの構築 | 日本電子計算(株) |
| | 17 アジア太平洋交流データベースの研究—プロトタイプ作成— | (株)西日本新聞社 |
| | 18 先端産業分野における専門用語の電子辞書データベース化の調査研究 | 科学技術情報研究所(株) |
| 技 術 | 19 書誌データベース用ダイナミックソーラス・エンジンの構築と自然言語検索システムへの応用 | (株)紀伊國屋書店 |
| | 20 知的ハイパーメディアを活用したデータベース構築に関する調査研究 | (株)新世代システムセンター |
| | 21 CD-ROM による光学材料データベースの構築 | (株)リアライズ社 |

目 次

| | | |
|-------|---------------------------|----|
| I | 各章の要約 | 1 |
| II | 記事データベースアクセス用 | |
| | パイロットシステムの構築と地方の情報化について | 7 |
| 1. | はじめに | 8 |
| 2. | パイロットシステムの構築体制 | 10 |
| 3. | パイロットシステム構築の背景 | 12 |
| 3.1 | 東北振興とデータベース | 12 |
| 3.2 | 新聞業界のデータベースへの取り組み状況 | 16 |
| 3.3 | 新聞記事データベースへの要望 | 18 |
| 3.4 | 東北地方のデータベースサービスの状況 | 19 |
| 3.5 | 「コミネット仙台」の現状 | 22 |
| 3.6 | データベース利用システム実現に伴う効果 | 26 |
| 4. | 新聞製作システムとデータベース | 27 |
| 4.1 | 新聞の製作フローとCTS | 27 |
| 4.2 | 記事データベースシステム概要 | 32 |
| 4.2.1 | データの流れとシステム構成 | 32 |
| 4.2.2 | キーワードと書誌情報項目 | 36 |
| 4.2.3 | データベース構築作業 | 38 |
| 4.2.4 | システムの特徴 | 42 |
| 5. | パイロットシステムの構築 | 43 |
| 5.1 | 記事データベースと「コミネット仙台」との接続の課題 | 43 |
| 5.1.1 | 配信システム構築の構想 | 43 |
| 5.1.2 | 「コミネット仙台」のシステム概要 | 43 |
| 5.1.3 | 接続のための課題 | 43 |
| 5.2 | パイロットシステム | 45 |
| 5.2.1 | 通信方式 | 45 |
| 5.2.2 | 通信制御装置の概要 | 45 |
| 5.2.3 | システムの接続とネットワーク・ソフト | 46 |
| 5.2.4 | パイロットシステムのテストと確認 | 46 |

| | |
|-------------------------|----|
| 6. 今後の展望 | 51 |
| 6.1 本システムの実現へ向けて | 51 |
| 6.2 システム発展の展望 | 52 |
| | |
| III 記事データベースアクセス用 | |
| パイロットシステムの構築に伴う技術資料 | 55 |
| 1. パシフィックシステム (CTS)構成図 | 56 |
| 2. CTSと記事データベースシステムの接続図 | 57 |
| 3. 記事データ蓄積処理フロー | 60 |
| 4. 通信制御装置構成図 | 61 |
| 5. 「コミネット仙台」ハードウェア構成図 | 64 |
| 6. アプリケーションシステム構成 | 65 |
| 7. 通信仕様 (X.25 プログラミング) | 68 |
| 7.1 X.25パケット交換とは | 68 |
| 7.2 コミネット仙台との接続 | 73 |
| 7.3 資源定義 | 75 |
| 7.4 アプリケーション・プログラミング | 76 |

I 各章の要約

[各章の要約]

記事データベースアクセス用

パイロットシステムの構築と地方の情報化について

1. はじめに

多様な可能性を秘めた情報化の波が、今地方で湧き起ころうとしている。

東北振興を掲げ、地方の情報化の進展を図る具体的方策のひとつとして、地域内で自らがデータベースを構築し、同時にまた、その情報を社会へ発信していく仕組みをつくっていくことがあげられる。

そこで本事業は、河北新報社（受託社）が独自に社内で利用することを主な目的として構築した記事データベースシステムについて、地域内パソコンネットワークを通じ外部のユーザーが利用できるよう、データベース利用システムを構築することを目指し、その実現を図り易くするためのパイロットシステムを構築した。具体的には、記事データベースと地域内パソコンネットワークという異なるシステム接続の課題を解決し、通信制御装置と標準ソフトの搭載、さらにアプリケーションソフトの開発を図ったものである。

2. パイロットシステムの構築体制

河北新報社情報局内に情報局長を中心に事務局を作り、外部スタッフに検討委員を依頼し、パイロットシステムの構築とそれに伴う課題の検討を行った。

（詳しくは本文参照のこと）

3. パイロットシステム構築の背景

3.1 東北振興とデータベース

東北には社会へ向けて発信する情報が、本当に少ないのだろうか。東北発信の情報で、全国、全世界から需要のある情報はないのだろうか。答えは「否」である。なぜなら、東北には世界的に注目される研究機関が多数存在している。例えば、東北大学における電気通信、金属材料、脳外科をはじめ秋田大学の鉾山、青森の雪の分野などいくらでも需要のある情報があるはずである。

しかし、残念なことに今までこれらの情報は、バラバラにしかも小出しに発信されることはあっても、東北が一体となって発信し、その情報を活用していくという体制がつくられていなかった。

今こそ、ここに東北が一体となって情報の新しい流通の仕組みをつくっていく必要がある。その試みの第一歩が、地域データベースの構築であり、それを活用しての情報の発信である。地域データベースを通して東北の情報が全国、全世界へと発

信されることによって、東北の存在が多面的に認識され、その結果として東北の振興が図られていく。

3.2 新聞業界のデータベースの取り組み状況

編集支援用の記事データベースは、東京大手紙のみならず、大手の地方新聞社も何らかの対応をしている。

しかしながら商用化への取り組みは、朝日、毎日、読売、日経等の東京大手紙を除けば、まだまだ不十分と言わざるをえない。

3.3 新聞記事データベースへの要望

1991年(平成3年)7月に受託社が仙台市内の企業、自治体を対象に受託社が構築すべき新聞記事データベースについて聞き取り調査を実施した結果、データベースへのニーズと期待はかなり高いものがあった。

そして、欲しい情報、機能としては、東北6県の地域ニュース、訃報、速報サービスなど前向きな意見が出されている。

3.4 東北地方のデータベースサービスの現況

東北地方において、データベースの一般的なディストリビュータとしては、「コミネット仙台」と各県の「中小企業情報センター」があげられる。

パイロットシステムを接続するネットワークとしては、多方面に会員を有している「コミネット仙台」が望ましいと思われる。

3.5 「コミネット仙台」の現状

「コミネット仙台」は、1986年(昭和61年)12月、日本初の第三セクターによるパソコンネットワークとして設立された。

サービスは、24時間運用で、アクセスポイントは全国に125カ所、月間のアクセス回数とその時間は、平均して約11,000回、1回約3分となっている。データベースについて1988年(昭和63年)のサービス開始時には、6種類のデータベースを提供するにすぎなかったが、現在では約4倍の種類に増え、利用回数もサービス開始当初の6倍、月間利用回数で1,300件を数えるところまで成長している。

3.6 データベース利用システム実現に伴う効果

地方から発信される情報についてその扱いは、全国ネットとローカルネットとは当然異なってくる。地方の情報は、地方の手によって発信されなければ、地方の実情は伝わらないと言えよう。その意味で、このパイロットシステムが第一歩となり、本システムが構築され、そしてデータベース利用システムが実現されることになれば、地方の情報の全国への発信が積極的に行われることになり、東北の情報

化および活性化を促進する効果をもたらす。

同時に、利用者の情報に対する取組姿勢が積極的なものに変化し、次のデータベースへのニーズを生み出す効果も期待できよう。

4. 新聞制作システムとデータベース

4.1 新聞の製作フローと CTS

新聞は、記者による取材から始まって新聞として製作、印刷、配達されて読者の手に渡る。この新聞製作のプロセスをさらに詳しくしてみると、入力システム、校正システム、組版システム、版下出力システム、製版システムおよび印刷、発送システムといったシステム構成の流れをたどって作業が進められている。このうち版下出力システムまでを CTS (Computerized Typesetting System) と呼んでいる。

4.2 記事データベースシステム概要

4.2.1 データの流れとシステム構成

データ蓄積に伴う基本的作業は、CTS からのデータの自動取り込みにはじまり、記事の電子的切り抜き、記事データと記事イメージの蓄積、不要なコマンドの除去、辞書の整備、そしてデータ項目の追加である。

システム構成として、主なハードウェアはホストマシン (IBM9121/320)、切り抜き端末、FAX サーバーおよび検索端末。ソフトウェアとしては、基本ソフト、画像処理ソフト、データベース処理ソフト、その他アプリケーションソフトがある。

4.2.2 キーワードと書誌情報項目

キーワードは、取り込まれた記事電文から自動的に抽出される。抽出の処理はキーワード抽出処理、ひらがな抽出処理、1文字削除処理、および複合語処理といった流れをとっている。これら各処理は、キーワード抽出辞書、ひらがな辞書、1文字辞書を用いることにより行われる。

また、本文中に存在しない項目をキーワードとしたい場合は、その項目をデータ項目と称して入力し、記事データに付加している。

なお、そのほかの辞書として、検索時に使用する同義語辞書を整備している。

4.2.3 データベース構築作業

実際のデータベースの構築作業は、基本的に記事の切り抜き作業と書誌情報の入力、そしてデータの蓄積などであり、これらの作業の繰り返しとなる。その他、関連した作業としては、辞書のメンテナンスなどがある。

4.2.4 システムの特徴

受託社のデータベースシステムの特徴としては、新聞製作システムからのデータの自動取り込み、電子的切り抜き方式の採用、紙面のデータベース化の範囲の広さ、記事イメージをそのまま FAX に出力できる等々があげられる。

5. パイロットシステムの構築

5.1 記事データベースと「コミネット仙台」との接続の課題

5.1.1 配信システム構築の構想

一般のユーザーが受託社の社外から利用できるようにするためには、まずデータを外へ出す仕組み（配信システム）を作る必要がある。具体的には、受託社のシステムとユーザーのシステムをつないでオンラインで利用する方法や、受託社からディストリビューターへ記事電文を送るといった方法がよく知られている。

5.1.2 「コミネット仙台」のシステム概要

「コミネット仙台」のシステムは、(株)富士通の A-60 と呼ばれるコンピューターを中心に、MSIA (Multiple Serial Interface Adapter) 6 セットのほか増設メモリ、ディスクなどで構成されている。

5.1.3 接続のための課題

両システムを接続しようとした場合、当面解決を要する技術的課題として、両システム間の通信方式、通信手順の決定、通信方式・手順に基づいてソフトウェアを開発するためのツールの開発などが考えられる。

5.2 パイロットシステム

上記のような接続の基本的な課題を短期間に一挙に解決することは困難である。そこで、まず通信上の主要な課題を解決して、パイロットとなるシステムを構築、そののち早期に本システムを構築することとし、以下のようなシステムを構築した。

5.2.1 通信方式 X.25

両社間の通信方式を、CCITT の X.25 とし、第 1 種パケット交換回線 1 回線を開設した。

5.2.2 通信制御装置の概要

採用した通信制御装置は、IBM の 3745/170 で、チャンネル・アダプターを 4 つ持ち、最大 112 回線と接続できる。設置したシステムは、19.2 Kbps で 16 回線に

対応可能の構成になっている。

5.2.3 システムの接続とネットワーク・ソフト

両システムの接続は、受託社のコンピューター IBM9121/320、通信制御装置 3745/170、DSU (Digital Service Unit)、NTT の回線 DDX-P、DSU、「コミネット仙台」のコンピューター A-60 といった構成になっている。そして、この構成上で基本ソフトとアプリケーションソフトが動作している。

5.2.4 パイロットシステムのテストと確認

パイロットシステムの構築に伴い「プロトコル接続確認テスト」を実施し、接続の確認を行った。

6. 今後の展望

6.1 本システムの実施へ向けて

パイロットシステムの目的は、本システムの実現を容易にすることにあり、早急に本システムを構築し、サービスを開始していくことが望まれる。

パイロットシステム構築後、本システムの実現へ向けては、検索のコマンド体系統一の必要性や「コミネット仙台」のユーザーへのファクシミリサービスの提供などいくつかの課題の解決が必要とされている。

6.2 システム発展の展望

当面の目標である本システムが実現し外部へ配信する仕組みができたとして、受託社のデータベースには、アクセス可能なネットワークをどのように拡大していくか、サービスの質および量をどのように充実させていくかなどという基本的な課題がある。これらを今後どのように考えていくか、システムの発展性と大きく係わってくる。

II 記事データベースアクセス用パイロット システムの構築と地方の情報化について

1. はじめに

多様な可能性を秘めた情報化の波が、今地方で湧き起ころうとしている。政治、経済、文化等の様々な面で東京圏に多くの機能が集中し、適切な機能分担が阻害されている今日の状況に、対応して求められつつある多極分散型国土形成の動きである。この多極分散の動向を地方でとらえ、実現していくことができるかどうかは、地方において、情報システムや通信ネットワークなどの情報化の基盤がいかに整備されているかと大きくかわってくる。

東北地方においては、東北インテリジェント・コスモス構想が「学術・技術・情報機能の集積と高度化」を基本戦略に掲げ、この基本戦略を具体化するため種々の活動を展開している。

東北振興を掲げ、地方の情報化の進展を図る具体的方策のひとつとして、地域内で自らがデータベースを構築し、同時にまた、その情報を社会へ発信していく仕組みをつくっていくことがあげられる。

そこで本事業は、河北新報社（受託社）が独自に社内で利用することを主な目的として構築した記事データベースシステムについて、地域内パソコンネットワークを通じ外部のユーザーが利用できるよう、データベース利用システムを構築することを目指し、その実現を図り易くするためのパイロットシステムを構築した。具体的には、記事データベースと地域内パソコンネットワークという異なるシステム接続の課題を解決し、通信制御装置と標準ソフトの搭載、更にアプリケーションソフトの開発を図ったものである。

今後これを更に発展させることにより、平成4年8月の記事データベースシステム社内サービス開始に合わせて、パソコンネットワークを通じての社外サービスが実現されるものと期待される。

今日、新聞記事のデータベース化は、多くの新聞社によって取り組まれている。しかし、マーケットの先行きの不透明さや費用面、技術面の困難さがあるため、“十分進展している”とは言えない状況にある（3.2参照）。そのような中で受託社のデータベース構築への取り組みには、地域の大きな期待が寄せられている。

なお、受託社が開発した記事データベースは、平成3年8月からデータを蓄積しており、主に次のような特徴をもっている（4.2参照）。

- ① 日常発行される新聞記事のほぼ全量を蓄積の対象にしている。
- ② 新聞電子組版システム（CTS）で作成された紙面を電子的に切り抜くことによって記事データを蓄積している。
- ③ 必要に応じて切り抜きイメージをFAXで提供できる。
- ④ 検索システムとして、“漢字・かな検索方法”を採用している。

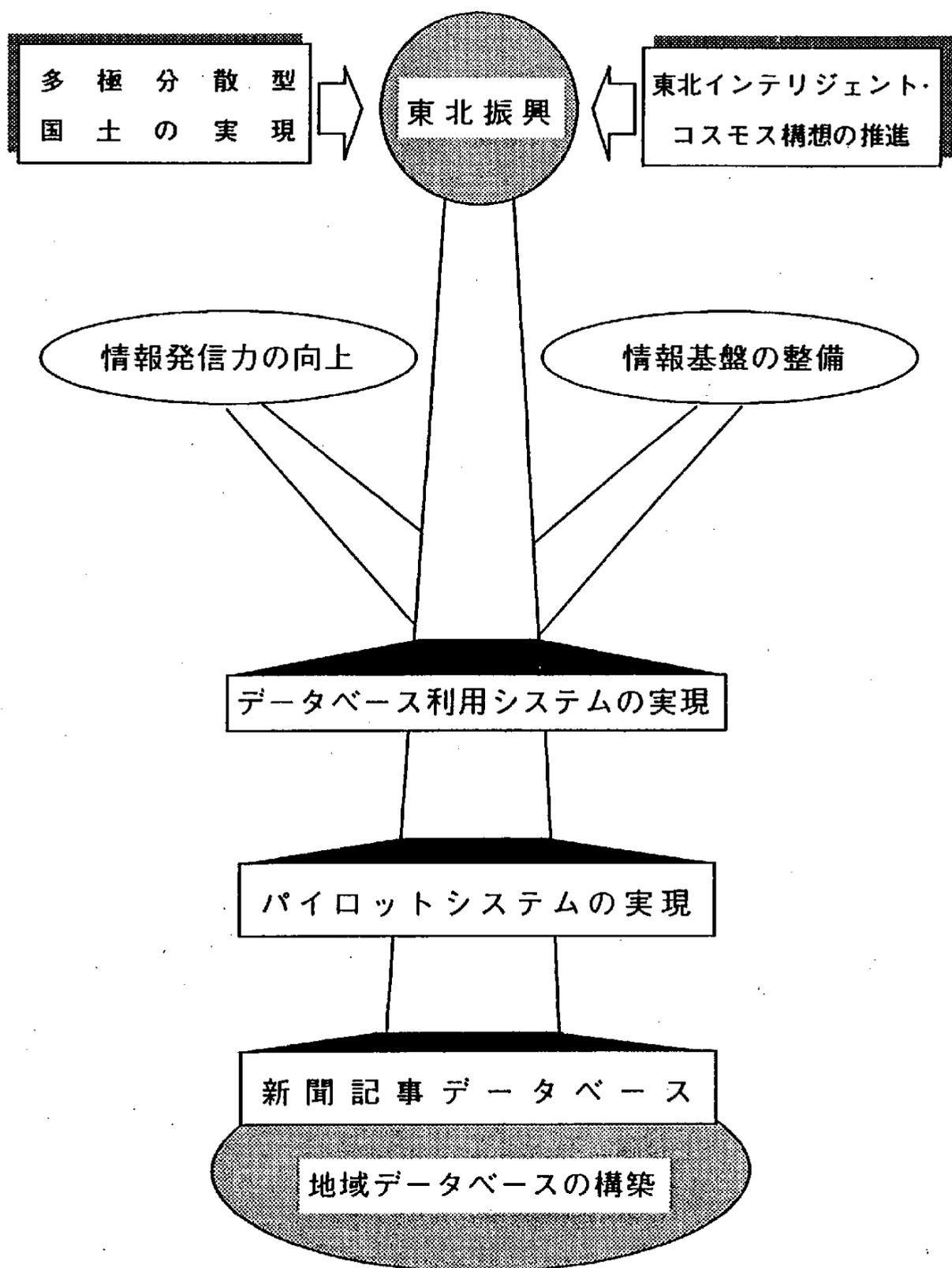


図1-1 記事データベースアクセス用パイロットシステム構築の意義

2. パイロットシステムの構築体制

河北新報社情報局内に情報局長を中心に事務局を作り、外部スタッフに検討委員を依頼し、パイロットシステムの構築とそれに伴う課題の検討を行った。委員会の構成メンバーは次の通りである（氏名表示は、五十音順による）。

2.1 記事データベースアクセス用パイロットシステム構築委員会

| | | |
|-----|-------|-----------------|
| 委員長 | 宮崎 正俊 | 東北大学 |
| 委員 | 稲葉 輝雄 | (株)東北システムズ・サポート |
| | 大澤 隆夫 | (株)コミネット仙台 |
| | 大山 肇 | 日本アイ・ピー・エム(株) |
| | 奥山恵美子 | 仙台市 |
| | 田宮 利和 | (株)富士通 |
| | 田村 政勝 | NTT データ通信(株) |

| | |
|------|--|
| 担当業務 | <ul style="list-style-type: none">・配信システムの仕様の決定・システム接続に伴う課題の解決・パイロットシステムの役割の評価・地域情報化のあり方の検討 |
|------|--|

2.2 記事データベースアクセス用パイロットシステム構築専門委員会

| | | |
|------|-------|----------------------|
| 専門委員 | 伊藤 康彦 | (株)コミネット仙台 |
| | 大宮 孝司 | 河北新報社 |
| | 海藤 恒久 | (株)富士通東北システムエンジニアリング |
| | 早川 徹 | 河北新報社 |
| | 平野 覚 | (株)東北システムズ・サポート |
| | 古内 博勝 | NTT データ通信(株) |
| | 星 光義 | NTT データ通信(株) |
| | 町田 章 | 日本アイ・ピー・エム(株) |

担当業務： 記事データベースアクセス用パイロットシステム構築委員会（本委員会）の委託に基づき、パイロットシステム構築に伴う実作業を行い、その結果について本委員会に報告する。

2.3 事務局

| | | |
|------|-------|-------|
| 事務局長 | 杉本 安克 | 河北新報社 |
| | 後藤 義雄 | 河北新報社 |
| | 千田 利昭 | 河北新報社 |

担当業務： 委員会の運営並びに本事業の遂行に必要な一切の業務を行う。

3. パイロットシステム構築の背景

パイロットシステム構築の背景を東北振興とのかかわりから検討し、新聞業界のデータベースへの取り組み状況と記事データベースへの要望、東北地方のサービス状況、「コミネット仙台」の現状を調べ、データベース利用システム実現に伴う効果を明らかにしていく。

3.1 東北振興とデータベース

“情報があるところに人が集まり、人が集まると情報が膨らむ”と言われている。現在、我が国では、東京圏に都市機能が集中し、人口も情報も過度の集積をみせ、この言葉通り「東京対地方」の情報格差は人口比を大きく上回っている。データベースサービスの売上高は、その8割までが東京圏で占められており、東京圏への情報の集中とそこから発信される圧倒的な情報量を物語っている。

東京圏への一極集中は、地方を産業活動、社会活動、個人生活等様々な面において東京圏と結び付かなくては、単独では機能しえない状況に追い込みつつある。つまり、東京圏と地方の情報、教育、文化等の量的・質的格差や就業機会の不足等が、地方への定住化や地方の活性化を阻んでいる。

近年、北海道や中国、九州地方など、全国各地で地域情報化、とりわけ地域データベースに関する議論が活発化してきている。これは、地方が将来的に都市としての機能を維持し、均衡ある国土の発展に寄与していくための方法論の模索でもある。

「情報」について、何らかの対策をとらない限り、地方の発展は成り立たないと言えよう。

「南の福島が東京・関東圏の強い影響を受け、また北の青森も北海道圏との相互交流が深まりつつあるという東北の置かれた立場は、深刻」という声さえある。東北独自の文化を維持しつつ、将来へ向けて発展していくためには、今この時を捉え情報化への積極的な努力をしていく必要があるのではないだろうか。

そのためには、東北の特性を活かしたデータベースを構築し、その情報を全国へ、そして世界へと向けて発信していく仕組みをつくらなければならない。科学技術の発達と情報ニーズの高まりにより、社会全体の情報流通量は増加の一途をたどっている。しかし、東北の各県についてみた場合、消費情報量では平均的な位置にあるものの、供給情報量では平均をかなり下回る位置にある(図3-1、表3-1)。これは、東北として発信する情報が、非常に少ないことを意味している。

東北には社会へ向けて発信する情報が、本当に少ないのだろうか。東北発信の情報で、全国、全世界から需要のある情報はないのだろうか。答えは「否」である。なぜなら、東北には世界的に注目される研究機関が多数存在している。例えば、東北大学における電気通信、金属材料、脳外科をはじめ秋田大学の鉱山、青森の雪の分野などいくらかでも需要のある情報があるはずである。

しかし、残念なことに今までこれらの情報は、バラバラにしかも小出しに発信されることはあっても、東北が一体となって発信し、その情報を活用していくという体制がつかわれていなかった。その結果、東北の様々な情報は、東北において直接流通されるということではなく、東北の情報でありながら、東京圏から発信された情報として、入手するという「情報の逆輸入現象」と呼ばれる不思議な状況を引起している。今こそ、ここに東北が一体となって情報の新しい流通の仕組みをつくっていく必要がある。その試みの第一歩が、地域データベースの構築であり、それを活用しての情報の発信である。地域データベースを通して東北の情報が全国、全世界へと発信されることによって、東北の存在が多面的に認識され、その結果として東北の振興が図られていく。

(単位:10⁴ワード)

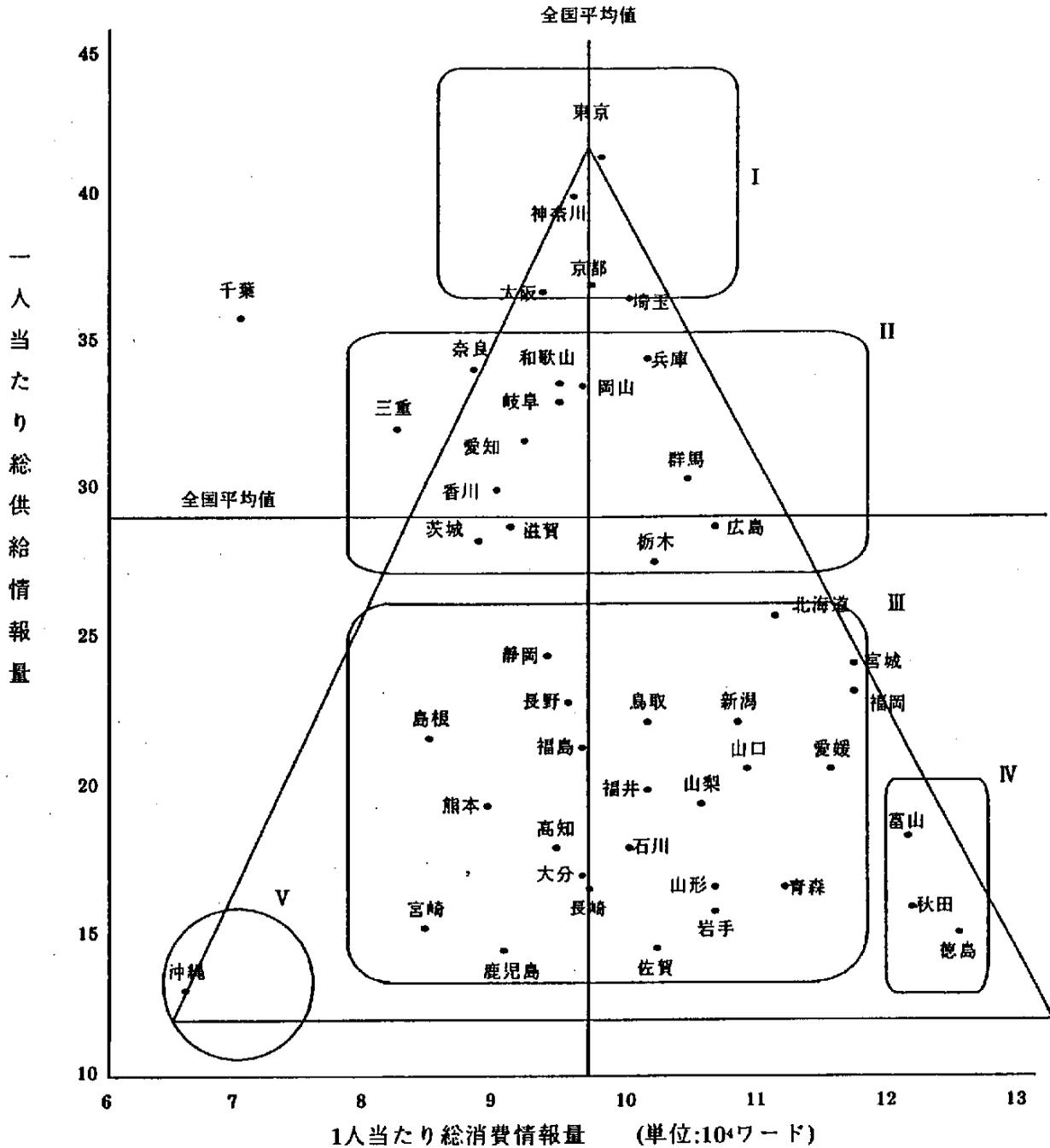


図3-1 県民1人当たりの総供給情報量と総消費情報量

(注) 情報とは、「情報流通センサス」(郵政省)の中において情報流の実態をとらえるため、あらゆるメディアによる情報流通の量を教えるの尺度で計算したものである。その計算方法を定めるに当たっては、次のとおりいくつかの前提を置いている。

- ① 対象は郵便(手紙、葉書)、電報、電話、データ通信、テレビジョン放送、ラジオ放送、新聞、書籍対話、学校教育、観劇等12種類とする。
- ② 各メディアの情報流通を、(a)言語、(b)音楽、(c)静止画、(d)動画の4つのパターンに分類し、その間に「換算比較」(例えばテレビジョン放送1分間の情報量は葉書何通分に相当するか)を設定するとともに、各メディアに共通の単位として日本語の1語を基礎としたワードという単位を設け、これによりすべての情報量を換算集計する。
- ③ 情報のもつ「意味」あるいは「価値」について計算の対象としない。
- ④ マス・メディア以外のメディアでは総供給情報量はすべて消費されるものとする。

表3-1 総供給情報量

| | 総供給情報量 | | | 構成比 (%) | | |
|-------|--------|--------|--------|---------|-------|-------|
| | 昭和50 | 昭和55 | 昭和60 | 昭和50 | 昭和55 | 昭和60 |
| 全国計 | 19,091 | 25,303 | 35,331 | 100.0 | 100.0 | 100.0 |
| 北海道 | 859 | 1,157 | 1,462 | 4.5 | 4.6 | 4.1 |
| 東北 | 1,225 | 1,616 | 2,368 | 6.4 | 6.4 | 6.7 |
| 青森 | 156 | 193 | 234 | 0.8 | 0.8 | 0.7 |
| 岩手 | 115 | 171 | 215 | 0.6 | 0.7 | 0.6 |
| 宮城 | 304 | 383 | 519 | 1.6 | 1.5 | 1.5 |
| 秋田 | 113 | 157 | 192 | 0.6 | 0.6 | 0.5 |
| 山形 | 110 | 165 | 196 | 0.6 | 0.7 | 0.6 |
| 福島 | 181 | 234 | 448 | 0.9 | 0.9 | 1.3 |
| 新潟 | 246 | 313 | 564 | 1.3 | 1.2 | 1.6 |
| 関東 | 7,749 | 10,292 | 14,223 | 40.6 | 40.7 | 40.3 |
| 東海 | 2,134 | 3,002 | 4,176 | 11.2 | 11.9 | 11.3 |
| 北陸 | 335 | 426 | 545 | 1.8 | 1.7 | 1.6 |
| 近畿 | 3,692 | 4,798 | 6,974 | 19.3 | 18.9 | 19.7 |
| 中国 | 1,112 | 1,412 | 2,087 | 5.8 | 5.6 | 5.9 |
| 四国 | 461 | 583 | 889 | 2.4 | 2.3 | 2.5 |
| 九州 | 1,435 | 1,902 | 2,466 | 7.5 | 7.5 | 7.0 |
| 沖縄 | 89 | 115 | 141 | 0.5 | 0.4 | 0.4 |
| | | | | | | |
| 三大都市圏 | 10,956 | 14,234 | 20,121 | 57.4 | 56.3 | 57.0 |
| 東京圏 | 6,397 | 8,400 | 11,673 | 34.5 | 33.2 | 33.1 |
| 名古屋圏 | 1,416 | 1,821 | 2,613 | 7.4 | 7.2 | 7.4 |
| 大阪圏 | 3,144 | 4,013 | 5,835 | 16.5 | 15.9 | 16.5 |
| 地方圏 | 8,135 | 11,069 | 15,210 | 42.6 | 43.7 | 43.0 |

(資料：郵政省「情報化の現況」)

3.2 新聞業界のデータベースへの取り組み状況

新聞社主要各社のデータベースへの取り組み状況は、表3-2の通りである。これを見ると、編集支援用の記事データベースについては、東京大手紙のみならず、大手の地方新聞社を含めて何かしらの対応がなされているようである。これは、CTSと呼ばれる新聞製作システム(4.1参照)の普及に伴い、系統的に記事データの自動蓄積が可能になったことによるものと考えられる。

しかしながらデータベースの商用化となると、その状況には厳しいものがある。朝日、毎日、読売、日経等の東京大手紙を除けば、商用化への取り組み状況は、まだまだ不十分と言わざるをえない。

このような中での本事業の取り組みは、情報化の立ち遅れている東北にインパクトを与えると同時に、他の地方新聞社に対しても様々な影響を及ぼすものと考えられる。

表3-2 主要各社のデータベース事例

| 新聞社名 | 編集支援データベース | | データベースの商用化 | |
|--------------------|------------|--------|---------------|--------|
| | 種類 | 稼働年月 | ネットワーク | 開始年月 |
| 朝日 | 記事、図書 | 85年11月 | 日経、G-Searchほか | 86年4月 |
| 毎日 | 記事、写説 | 88年4月 | 日経、G-Searchほか | 87年11月 |
| 読売 | 記事、写説 | 86年9月 | 日経、G-Searchほか | 87年4月 |
| 日経 ジャパン タイムス | 記事、イメージほか | | 日経、G-Search | 70年 |
| 日刊工業 | 記事 | 84年 | 日経 | 86年9月 |
| 日本農業 | 記事 | | 日経、G-Searchほか | 84年 |
| 共同通信 | 記事、写説ほか | 88年4月 | 日経 | 88年10月 |
| 北海道 | 記事 | 90年11月 | G-Search | 89年6月 |
| 上毛 | 見出し、人物 | 86年4月 | オーロラネット、日経 | 91年11月 |
| 山梨日日 | 見出し、写説 | 89年7月 | | |
| 静岡 | 記事、見出し | 81年1月 | 日経、共同 | 89年5月 |
| 中日 | 記事 | 89年10月 | MI-Base | 84年1月 |
| 新潟 | 見出し、人物 | 88年9月 | | |
| 北国 | 記事、人物ほか | 88年3月 | | |
| 京都 | 見出し、イメージ | 89年1月 | | |
| 中国 | 記事、図書ほか | 88年1月 | | |
| 四国 | 人物 | | | |
| 愛媛 | 見出し | 88年1月 | | |
| 高知 | 見出し | | | |
| 西日本 | 記事 | | 日経、共同 | 89年9月 |
| 熊本日日 | 記事 | 86年10月 | 共同 | 89年6月 |
| 南日本 | 記事、見出し | 85年10月 | 日経 | 88年10月 |

(新聞協会調べから作成)

(注)

1. G-Search=富士通、平和情報センターなどが創設したネットワーク。
2. オーロラネット=北海道新聞社が運営管理しているパソコン通信ネット。
3. MI-Base=NTTデータ通信のデータベースシステム。
4. 共同=共同連動データベースのこと、G-Searchを通じて配信。
5. 写説=写真説明のこと。

3.3 新聞記事データベースへの要望

1991年(平成3年)7月に受託社が仙台市内の企業、自治体を対象に受託社が構築すべき新聞記事データベースについて聞き取り調査を実施した結果、データベースへのニーズと期待はかなり高いものがあった。

そして、欲しい情報、機能としては、次のような前向きな意見が出されている。

(1) 欲しい情報

記事データベースとして欲しい情報の種類には、日常の紙面に掲載される記事のほとんど全てがあげられた。とりわけ次の項目についての要望が多い。

- | | |
|---------------|---------------------------------------|
| ① 東北6県の地域ニュース | ・各県版のニュース ・6県各紙(他社)の記事 |
| ② 訃報 | ・6県各紙の訃報 |
| ③ プラスアルファの情報 | ・企業情報 ・経済情報 ・選挙情報 ・観光・イベント情報 |
| ④ 古い記事 | ・過去の10大ニュース ・10年前、40年前の情報も欲しい |
| ⑤ ビジュアル情報 | ・統計資料等 |
| ⑥ 広告 | ・特売情報 ・広告プラスアルファ情報 |

(2) 欲しい機能

機能についての要望では、検索の容易性はもとより、単なる新聞の切り抜き情報の提供にとらわれない次のような各種サービスが求められている。

- | | |
|-------------------------|----------------------|
| ① 付加価値サービス | ・整理情報の提供 ・加工情報の提供 |
| ② 速報サービス | |
| ③ 紙面に掲載されない記事(没記事情報)の提供 | |

これらの要望事項をみると、新聞記事データベースの構築にあたって、即実行可能と判断できるものは少ないかもしれない。しかし、いずれの要望も将来的に実現されれば、利用者にとって非常に魅力的なデータベースとなるであろう。

(3) データベースの一層の発展のために

以上(1)、(2)でみてきたような意見のほかに、データベースについて次のような声があ

ることも事実である。

- ① 情報収集は、社内のネットワークで十分。
(本社から情報が流れてくる)
- ② 新聞の切り抜きで十分。
- ③ データを引き出しても再整理に手間がかかる。
- ④ 欲しい情報がかめない。
- ⑤ 検索端末の操作が複雑だ。
- ⑥ 新聞を読まないといわれる世代が、記事データベースを果たして使うのだろうか。
- ⑦ 料金が高い。新聞購読料程度にしてほしい。
- ⑧ 内容に見合うものであれば、多少高額でも使用する。

何らかの形でデータベースに携わる者としては、このような声を謙虚に受け止め、データベースの一層の発展のために真摯に取り組んでいく必要があるだろう。

3.4 東北地方のデータベースサービスの状況

東北地方において、不特定多数の利用を想定し構築（稼働、運営）されている主なデータベースは、表3-3の通りである。これをみてわかるように、データベースの一般的なディストリビュータとしては、「コミネット仙台」と各県の「中小企業情報センター」があげられる。

本事業においてパイロットシステムを構築する場合、将来的に記事データベースの想定利用者は、一般から企業、大学関係者まで多方面に及ぶものと考えられる。このようなことから、パイロットシステムにおいて接続するネットワークとして、都市のコミュニケーションネットとして存在している「コミネット仙台」が望ましいと思われる。

表3-3 東北地方で構築されているデータベースサービス一覧

| 構築・運営主体(データベース名) | 概要 | 業態 |
|---|---|-----|
| 東北大学 金属材料研究所論文データベース | ・コンピュータによる材料設計の基礎資料となる種々の情報を集大成した材料データベース | P |
| 岩手県工業試験場 | ・新技術事業団が作成している新技術情報のデータベース 「ハイテクネットとうほく」(パソコン通信)のメニューとして提供 | D |
| (株)コミネット仙台 | ・商工名鑑、せんだいくらしのガイド等の地域データベース ・各新聞、企業信用等々の一般商用データベース | D |
| 宮城県中小企業情報センター MINES (Miyagi Information Network & Exchange System) | ・文献、企業情報、人材情報、統計情報等の地域固有データベース ・外部データベースサービス(SMIRSなど) | P・D |
| 青森県中小企業情報センター 秋田県中小企業情報センター 岩手県中小企業情報センター 福島県中小企業情報センター 山形県中小企業情報センター | ・SMIRSの提供および各県の文献情報、企業情報、人材情報など独自データベースを提供 | P・D |
| 宮城県 宮城県教育情報システム | ・能力開発情報に関する約10万件のデータベースを持つ中央能力開発協会と県職業開発協会をオンラインで接続し、人材、教育に関する情報を提供 | D |
| 宮城県 行政資料検索システム 総合統計情報提供システム | ・県が所有する行政資料の保有状況を提供 ・統計データの提供 | P |

| 構築・運営主体(データベース名) | 概 要 | 業 態 |
|---------------------------------------|---|-----|
| 宮城県 消費者生活情報ネットワークシステム (PIO-NET) | ・国民生活センターに設置されたホストコンピュータと接続し、情報の提供 (相談情報、危害情報、商品テスト情報、判例情報、文献情報など) | |
| 福島テレメディアサービス | ・郡山商工会議所調査による会員事業所のデータベースの提供 | P・D |
| 青森県立農業経営研究所 農業雑誌目次データベース | ・農業雑誌の目次を県の関係機関に提供 | D |
| 山形県立農業試験場 | ・アメダス・メッシュ気候データ、県内のアメダス観測データ、メッシュ気象値・実況値の情報を県の農業関係機関に提供 ・水稻ササニシキの成育データベースとして、農業試験場4箇所のササニシキの成育データを収録 | P |

P:プロデューサー D:ディストリビュータ

3.5 「コミネット仙台」の現状

経済の24時間化が進行するなかで、昼夜を問わず動き続ける都市のコミュニケーション手段のひとつとして、時間や空間の制約を越え、参加者の社会的属性にもとらわれない開かれた交流の場としてのパソコン通信が脚光を浴びてきている。

「コミネット仙台」は、仙台市がテレトピア構想の指定を受けたことを契機に、1986年(昭和61年)12月、日本初の第三セクターによるパソコンネットワークとして設立された。東北の中核都市としての情報インフラストラクチャーを地域あげて構築しようとの主旨から、株主、役員それぞれに地域を代表する機関、リーダーが参加している。

「コミネット仙台」のネットワークには、東北をエリアとする国等の出先機関、地元マスコミ、主力金融機関、宮城県、仙台市そしてそれらの管轄する各種機関、文化施設、更に企業、市民が参加している。

サービスは、24時間運用で、回線数が24。アクセスポイントは全国に125カ所、法人会員と個人会員とはおおよそ半々の割合である。月間アクセス回数とその時間は、平均して約11,000回、1回約3分となっている。

「コミネット仙台」が提供するデータベースサービスには、地域データベースと一般商用データベースがある。地域データベースには商工名鑑データベース、せんだいくらしのガイド、食い道楽百科等がある(表3-4)。また、一般商用データベースとしては、各新聞社記事、企業信用、人物情報、図書情報サービス等がある(表3-5)。

1988年昭和63年のデータベースサービス開始時には、6種類のデータベースを提供するにすぎなかったが、現在では約4倍の種類にふえている。そして、当然のことながらデータベースの利用回数もサービス開始当初の6倍、月間利用回数で1,300件を数えるところまで成長している(図3-2)。

一般商用データベースの利用分野としては、各マスコミ、企業信用情報、雑誌等の利用率が高くなっている。また、地域データベースとしては、無料サービスということもあるかもしれないが、商工名鑑の利用率が高い。

「コミネット仙台」の会員数は、現在約3,000名を越え、毎年700名程度の新規会員が見込まれている。

以上述べた「コミネット仙台」の現状から、受託社の記事データベースを「コミネット仙台」を通じて配信することは、以下のようなメリットがあるものと考えられる。

- ① 利用者にとって新たな端末機を設置する負担がない。
- ② 「コミネット仙台」で既にサービスされている地域情報（商工名鑑、生活ガイド等々）と合わせて、更なる地域の情報集積の高度化が図られる。
- ③ 将来的に、地域独自の情報加工の道が拓かれる可能性がある。
- ④ 「コミネット仙台」には、各種行政機関が会員として加入しているので、今後公共施設等（市民センターなど）への導入が見込まれる。

表3-4 「コミネット仙台」が提供している地域データベース

| | |
|---|---|
| 商工名鑑データベース 農業園芸データベース 新技術情報データベース | せんだいくらしのガイド コンベンション・宿泊施設案内 食い道楽百科データベース |
|---|---|

表3-5 「コミネット仙台」が提供している商用データベース

| | |
|---|--|
| マーケットサーチ 東京商工リサーチ企業情報 東洋経済企業情報 タイトルサーチ テクノサーチ 読売新聞記事データベース 朝日新聞記事データベース 日経新聞記事データベース 毎日新聞記事データベース 日刊工業記事データベース MANAGEMENT(経営実務情報) 図書情報データベース | アジア速報サービス アジアビジネス情報 産業統計データベース CD新譜速報 音楽CD総カタログ PRIZE(賞・受賞情報) BOOK-MONTHLY 週刊誌AERA記事 ジェトロエース 雑誌記事データベース WHO(人物・人材情報) 帝国データバンク企業情報 企業ニュース情報データベース |
|---|--|

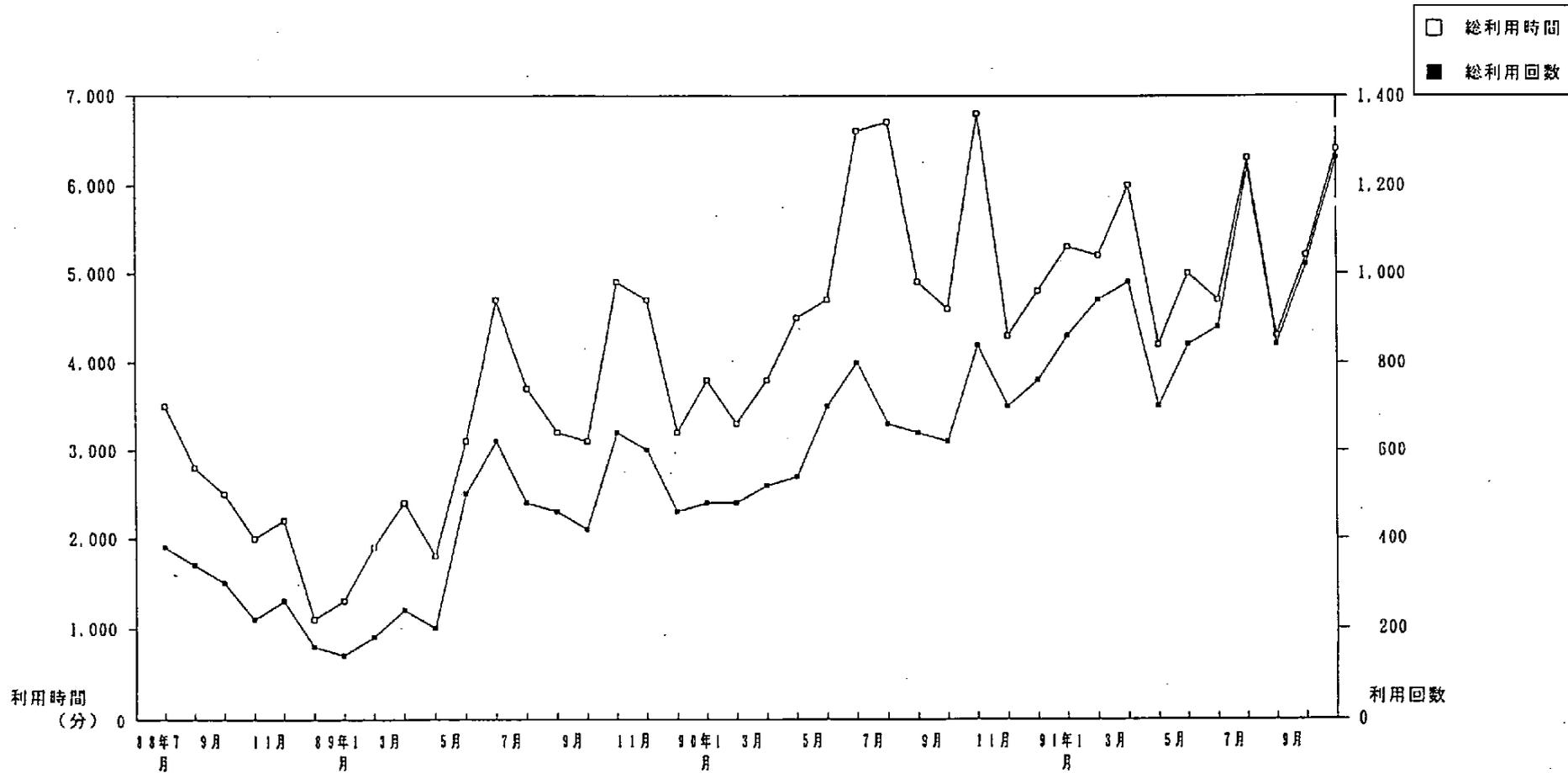


図3-2 「コミネット仙台」のデータベース利用状況の推移 (資料: 「コミネット仙台」提供)

3.6 データベース利用システム実現に伴う効果

新聞記事データベースの情報源の特性として、

- ① 活字情報が中心
- ② (When, Where, Who, What, Why, How, Howmuch)の表現に優れ、物事の経過がよくわかる
- ③ データの収録件数が膨大
- ④ あらゆる分野の情報が混在している

などの項目が一般的にあげられている。これらの項目から新聞記事データベースは、体系的な情報を幅広く求めるためのトータルガイダンスとして活用していくことができよう。

しかしながら、そこから得られる情報つまり検索結果は、利用するデータベースによって当然異なってくる。例えば、「日本経済新聞記事情報」では、企業、業界情報に高いヒット率を示し、朝日、毎日、読売の新聞記事情報では、「住民運動」や「市民運動」など生活情報よりのキーワードで高いヒット率を示している。新聞記事は、基本的に客観的立場に基づいて製作されるものだが、実際には新聞各社の方針や付加事業などの影響が少なからず出てくる。

従って、地方から発信される情報についてその扱いは、全国ネットとローカルネットとでは当然異なってくる。地方から全国に向けて発信したい情報であっても、全国ネットでは無視される場合がある。地方の情報は、地方の手によって発信されなければ、地方の実情は伝わらないと言えよう。その意味で、このパイロットシステムが第一歩となり、本システムが構築され、そしてデータベース利用システムが実現されることになれば、地方の情報の全国への発信が積極的に行われることになり、東北の情報化及び活性化を促進する効果をもたらす。

同時に、地方の身近な情報がデータベースとして提供されることにより、データベースの利用もまた身近なことになる。その結果、ユーザーの情報に対する取組姿勢が積極的になり、次のデータベースへのニーズを生み出す効果も期待できよう。

4. 新聞製作システムとデータベース

当該データベースの特色の一つは、CTS と呼ばれている新聞製作システムから直接記事電文を取り出して、それを電子的に切り抜いてデータベースを構築する点にある。そこで、まず新聞製作とデータベースとのかかわりを述べ、そのうえで、より直接的なデータベースの構築の概要にふれる。

なお、詳しくは技術資料を参照されたい。

4.1 新聞の製作フローと CTS

新聞は、記者による取材から始まって新聞として製作、印刷、配達されて読者の手に渡る。この新聞製作のプロセスをさらに詳しくみると、入力システム、校正システム、組版システム、版下出力システム、製版システムおよび印刷、発送システムといったシステム構成の流れをたどって作業が進められている。(図4-1)

このシステム構成のうち、入力から版下出力までのシステムでは、コンピュータが中心的な役割を果たしている。そこで、新聞業界では、版下出力までを CTS (Computerized Typesetting System) と呼んでおり、かなり一般的な用語として用いている。

なお、受託社では、図4-2のシステム全体を“パシフィック”という愛称を用いているが、狭義には、CTS 部分の愛称としても用いている。

CTS を構成している各システムは、およそ以下のような機能と仕組みになっている。

(1) 入力システム

ここでいう入力システムとは、漢字キーボードやワープロなどによって、記者の書いた原稿をコンピュータへ入力するシステムのことをいう。1文字を2バイトの機械可読の符号に変換する事が主な機能で、日本の新聞社の多くは、共同通信社はじめ加盟社が1977年に定めた“CO77”と呼ばれる符号体系を採用している。

(2) 集配信システム

入力した記事電文は、集配信システムと呼ばれるシステムへ集まる。このシステムは、今日の新聞社ではごく普通に見られるシステムで、自社の記事のほか友好社や通信社が送信した記事電文も、回線を通じてオンラインで入力する。なお受託社の集配信システムは、東芝製のコンピューター (TOSBAC 7/70E) が中心になって構築されている。

また、現在の作業は、入力に先立って編集のデスクないしは整理記者が、価値判断、掲載の可否、掲載紙面の指定、見出し付けなどを行い、オペレーターがその指示に従って、掲載面名、見出し、記事本文、写真説明などを漢字キーボードによって入力する方法がとられている。

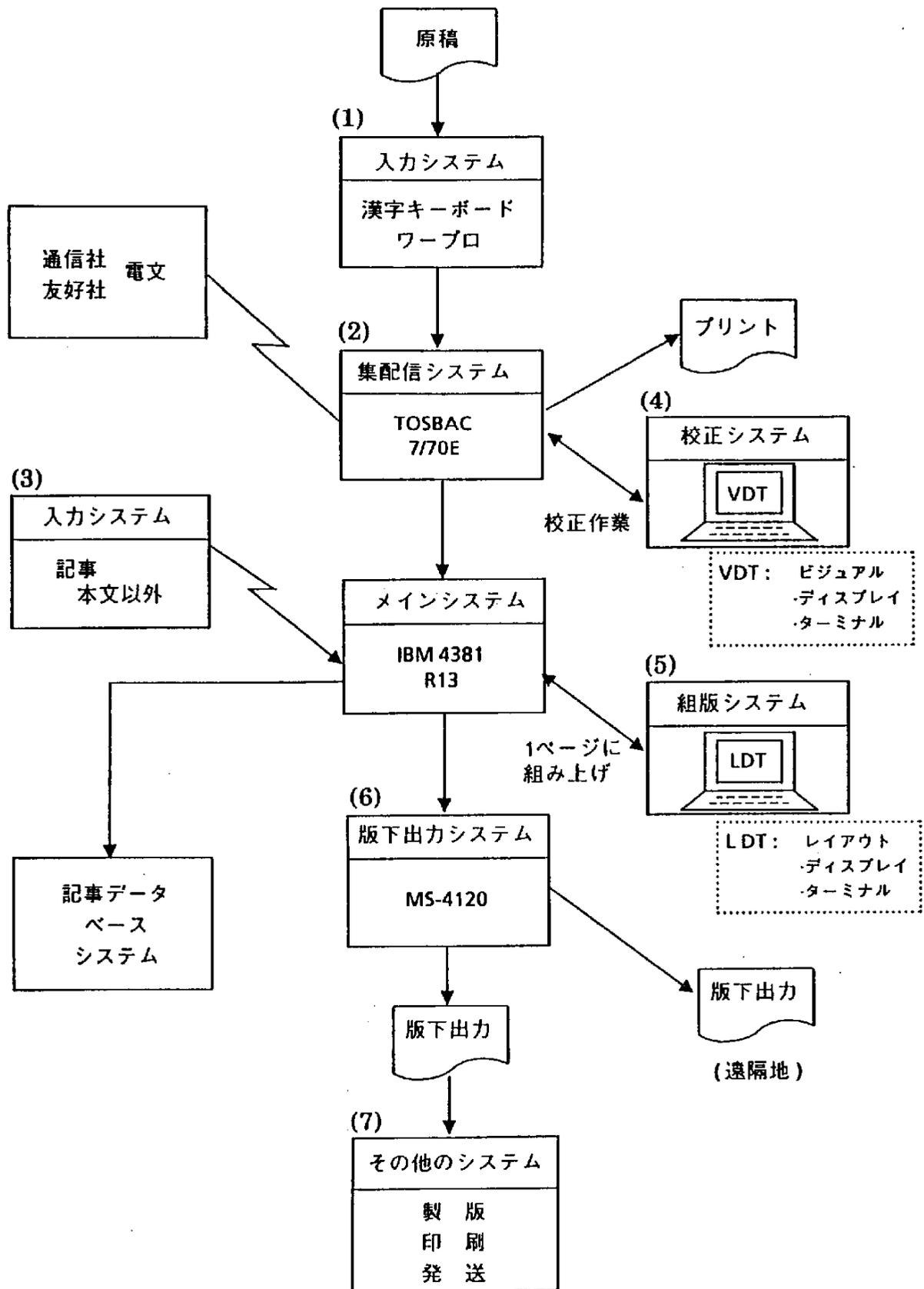


図 4-1 新聞の製作フローとCTS

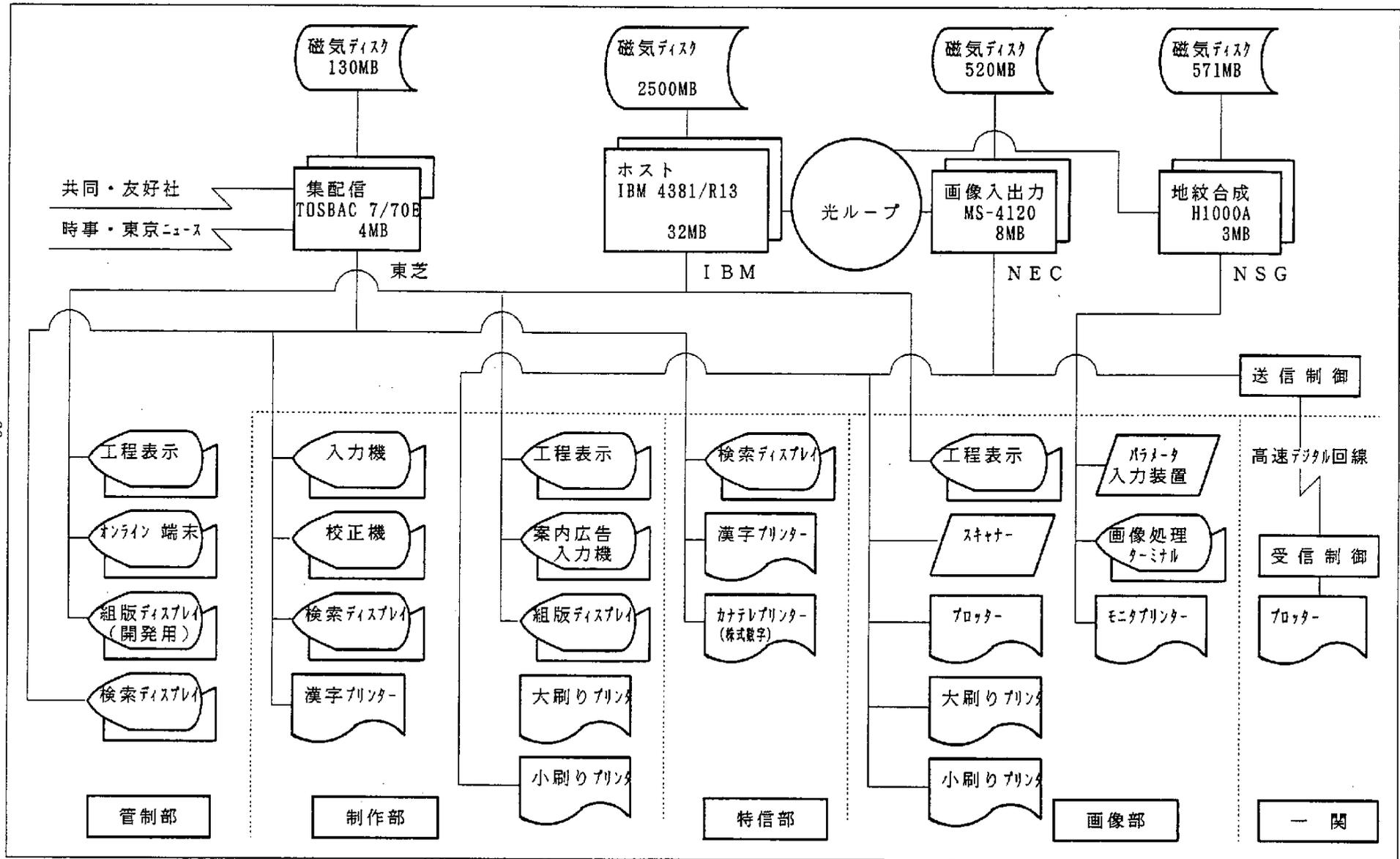


図4-2 パシフィック関連機器構成図

(3) その他の入力

紙面を構成している記事本文以外の要素、たとえば写真や図、地紋見出し、広告などは、記事本文の入力方法とやや違う流れをとっている。現状では、写真、図、広告は別に製作して、スキャナーによって入力しており、地紋見出しは、別にサブ・システムで作成してオンラインで全体のシステムとつないでいる。

※地紋見出し：地紋（網点や直線などの模様や図柄など）に見出し文字を重ね合わせたもの

(4) 校正システム

集配信システムは、文字通り信号となった記事を集信、配信する機能を持つが、そのほかにも入力した記事電文の内容をプリントとして出力する機能や、キーボードの付いた VDT 端末 (Visual Display Terminal) へ記事内を表示するとともに、その操作によって記事電文を校正するといった機能を持つのが普通である。

受託社では、記事内容のプリントを出力し、それを校閲記者が校閲、朱入れをしてオペレータへ回し、オペレータは VDT 端末を用いて集配信システムと会話する形で記事を訂正し、掲載する面などを指定して組版システムへ送り込んでいる。

(5) 組版システム

集配信システムから組版システムへ送り込まれた記事は、社説などに見られるような箱組み、見出し組みなど体裁が整えられたうえ新聞 1 ページの大きさに組み込まれる。

受託社の場合、組版システムの中心になっているのは、日本 IBM 社のコンピュータ (IBM 4381/R13) で、JPS-N と呼ばれるソフトウェアがベースになっている。これを JPS ホストとも呼んでいる。

このホスト・システムへ一時蓄えられた記事電文を LDT (Layout Display Terminal) と呼ばれる端末を操作することによって新聞 1 ページ大の大きさに記事を組み上げている。

電子的に組み上がった紙面は、さらに訂正のうえ、一定の時刻（新聞社では、一般に降版時刻といっている）になると出力指令のコマンドを打つことによって終了する。この出力指令により、1 ページ大の版下が出力する。同時に、同じ紙面内容がデータベース用のホストシステム (IBM9121/320) へも送られる仕組みになっている。

なお、この段階までの記事電文は、内部的にはいわゆるコード情報で、各社それぞれ少しずつ異なったコード体系を採用しており、受託社の場合は先の CO77 および IBM 社のコード体系が用いられている。

(6) 版下出力システム

ホストシステムからの出力指令により、文字は文字ファイルから、写真、図、地紋見出し、広告などはそれぞれのファイルからデータが引き出されて集められ、新聞1ページの大きさの版下の形で出力する。

受託社の場合、日本電気製のコンピューター (MS-4120) を中心とした出力システムで処理された紙面を、プロッターにより1ページ大のフィルムとして出力している。

なお、社内のシステムへの出力と同時に遠隔地 (一関市) に配置してあるプロッターへも出力している。

(7) その他のシステム

受託社では、出力したフィルムを用いて製版 (亜鉛版) し、オフセット輪転機で印刷、発送システムを経由して紙面を読者へ届ける、という流れをとっている。以下、データベースと直接の関連はないと思われるので省略する。

4.2 記事データベースシステム概要

(詳細は技術編参照)

4.2.1 データの流れとシステム構成

(1) データベース構築の基本的な流れ

受託社は、平成元年12月、基本的な構想を固め、2年4月から開発に着手、3年8月からデータの蓄積をはじめたが、そこにおける記事データベース構築作業の基本的な考え方は、およそ以下の通りである。(図4-3)

①データ取り込みの自動化

データは、前述のように、CTSにおける版下出力と同時に、同じ紙面、記事電文をデータベースのホストシステムへ取り込んでいる。

②切り抜き作業の電子化

必要な記事を切り抜く。

現在、日本経済新聞社で全自動的に切り抜く方法が実用化しているが、受託社では、自動化によって必要な記事を100%切り抜くことはまだ無理と判断、代わってパソコンを用いて、簡単に電子的切り抜きを行う方法を開発、実用化した。

③データベースとイメージ蓄積

切り抜いた記事を、記事データベースを構築するためのデータと、切り抜きイメージのままの記事の双方を蓄積する。

切り抜いたイメージ状に蓄積した記事は、後にイメージのままパソコンなどの端末で見たり、プリンタやFAXから取り出して見ることができるようにした。また、FAXとしては当面G3タイプに対応するとともに、将来普及が予想されるG4タイプにも対応できることを考慮した。

④不要なコマンド除去の自動化

データベースを効率的に構築するため、ソフトウェアを開発、データベースシステムへ入力した版下用の記事電文から不要なコマンド類を取り除く作業を完全自動化した。

⑤フリーキーワード方式の採用と辞書の整備

基本的には、記事電文から、自動的にキーワードを拾い出してデータベースを構築するといういわゆるフリーキーワード方式を採用、そのための辞書を整備している。

⑥データへのデータ項目の追加

フリーキーワード方式の弱点(キーワードが文章の内部に限られる)を補って、データを検索する際に便利なよう、シソーラスに代わるものとしてデータに“分類”などの付加情報を与えている。

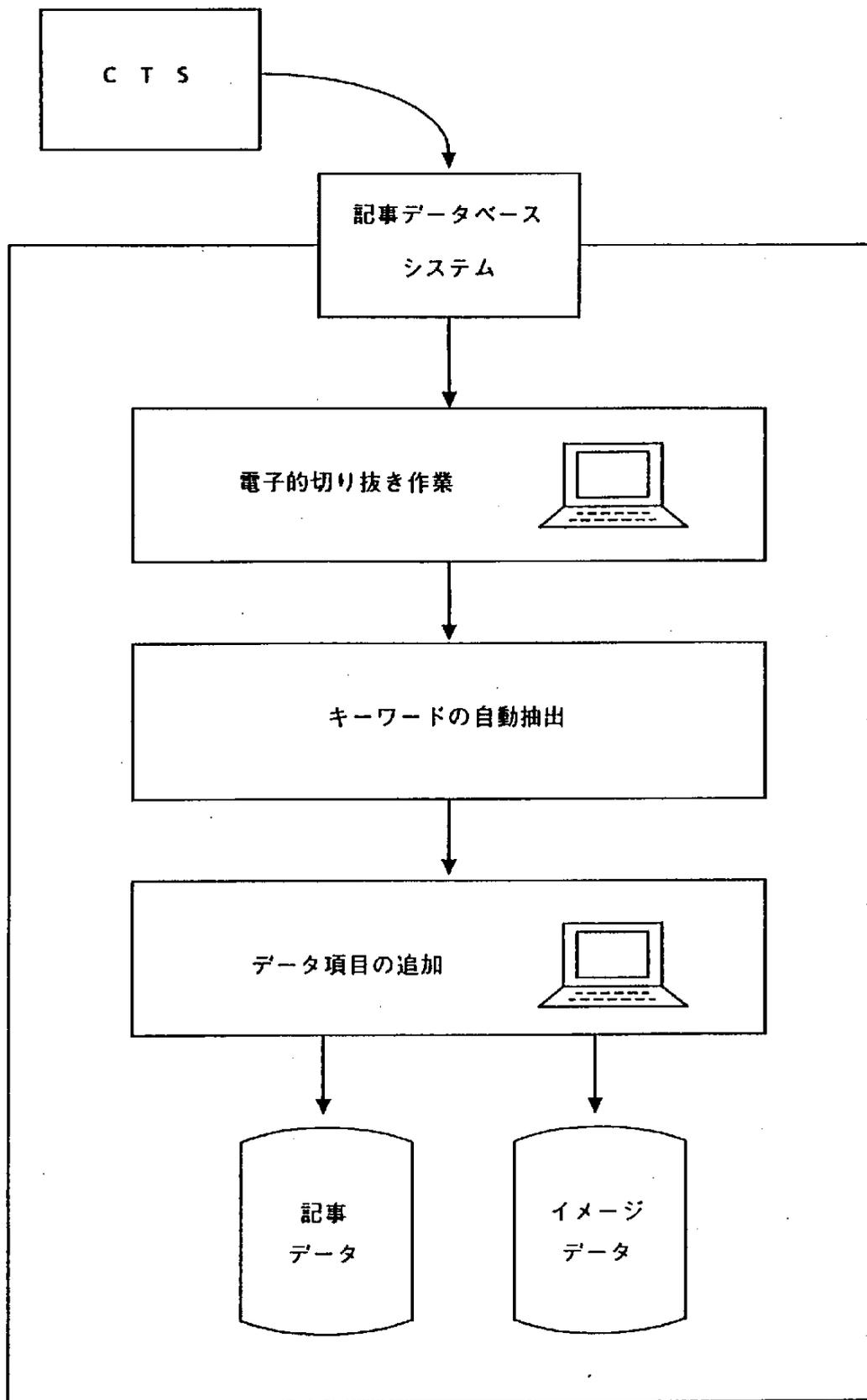


図 4-3 記事データベース構築フロー

(2) システム構成概要 (ハードウェア)

上記のような基本的な流れを実現するためデータベースとして整備したハードウェア構成は、ホストマシン、端末制御装置、切り抜き端末、FAX サーバーおよび検索端末から成っている。(図 4-4)

導入を決めたホストマシンはコンピューター IBM9121/320、記事切り抜き用端末はパソコン IBM5571-VIB、検索用の端末は同じくパソコン IBM5530-T03 が主な機器である。(技術編 2 節 CTS と記事データベースシステムの接続図参照)

これらの機器は、平成 2 年暮れ受託社内へ搬入、3 年 1 月、3 月の新聞休刊日に CTS 機器と接続、その後、記事蓄積のソフトウェアの開発を終えて 8 月から河北新報に掲載された記事を蓄積している。

(3) 基本ソフトウェア

導入したソフトウェアはおよそ以下の通りである。

①基本ソフト

基本オペレーティングシステム、データや装置の管理機能、開発者との対話処理機能など。

②画像処理ソフト

新聞記事イメージをコントロールする。

③データ・ベース処理ソフト

データベース制御のソフトウェア JAIRS/CICS(※)。

※ Japanese Information Retrieval System / Customer Information Control System)

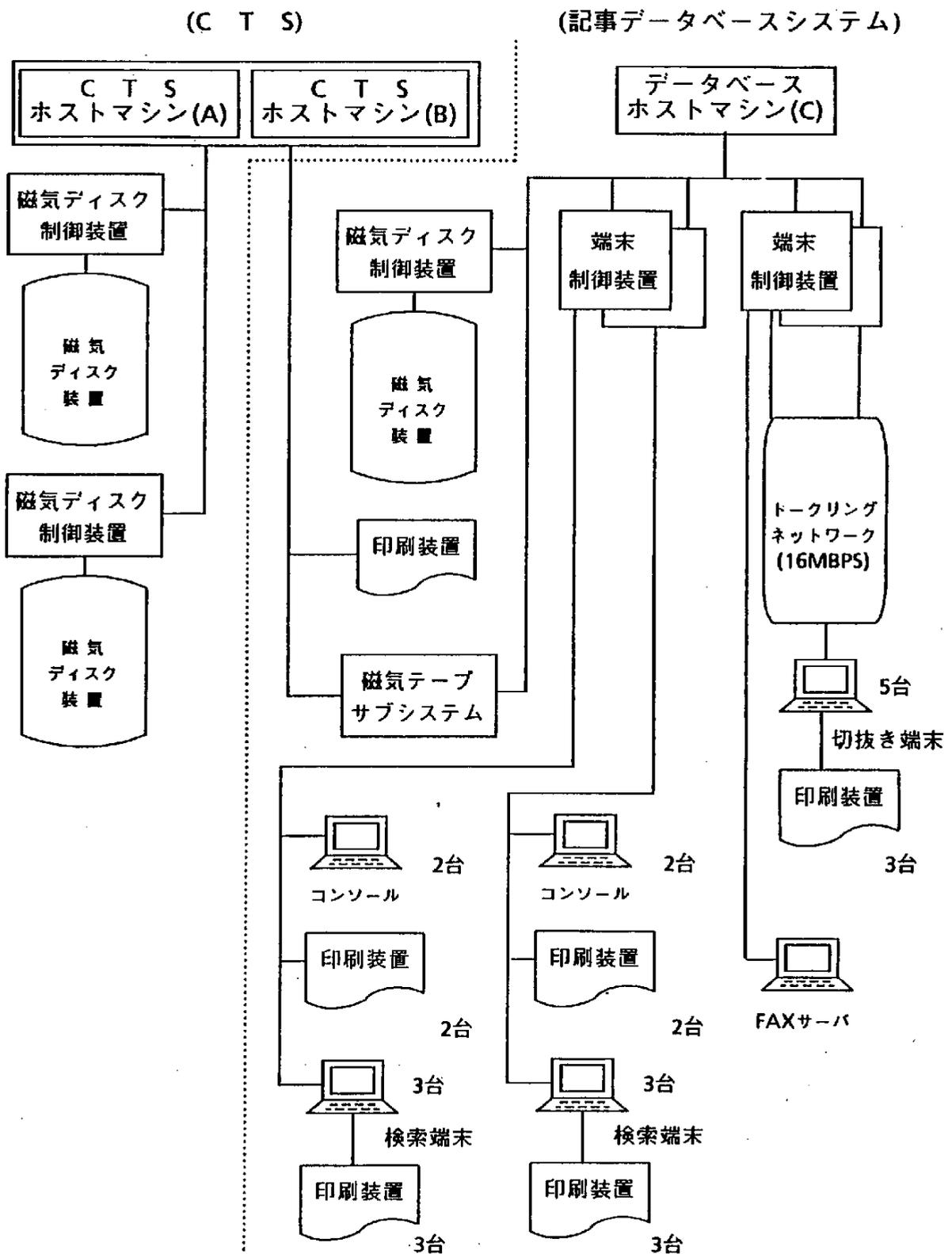
④プログラミング言語

アプリケーションソフト開発用の PL/I およびアセンブラー言語処理。

⑤開発したソフトの例

受託社は、EXT/SCF (EXternal/System Composing File) と呼んでいるオリジナル・ソフトを開発した。SCF とは、CTS の組み版ファイルのことで、このソフトを“組み版ばらし”と呼んでいる。

CTS からデータベースシステムへ流れてくる記事電文は、新聞紙面版下用の電文であり、1 文字 1 文字について、文字としての符号 (コード) のほかに、その文字の紙面上における位置情報、文字の大きさ、書体など各種情報が付け加えられている。それらはデータベース構築上からは不要な情報であり、それらをこのソフトで取り除いている。



◇ CTS、記事データベース両システム間でのデータ転送が必要なことから、ホストマシン(A)、(B)、(C)はどの磁気ディスク制御装置にもアクセスできる。

図 4-4 CTSと記事データベースシステム

4.2.2 キーワードと書誌情報項目

(1) キーワードの抽出と辞書

キーワードは、EXT/SCF を経由した記事電文から、自動的に抽出する。

抽出の処理は、キーワード抽出処理、ひらがな抽出処理、1文字削除処理および複合語処理といった流れをとっており、各処理は、それぞれ辞書を参照してワードを拾うという作業をしている。(図4-5)

なお、キーワード抽出辞書および複合語処理は、データベース処理ソフトJAIRSに標準として装備されている。他の辞書はユーザーが作成、メンテナンスを行う。

各辞書の内容、役割りは、おおよそ以下の通りである。

①キーワード抽出辞書

入力文書から単語を切り出すために参照される辞書で、文中の語と辞書中の語を比較し、最長で一致した語を一単語とみなしている。

キーワードの最小単位となり、この辞書の単語をもとに複合語処理され、抽出キーワード数の減少にもつながる。ほとんどの単語は提供される。

②ひらがな辞書

ひらがな、ひらがな混じりの語を抽出するために登録する。ここで登録されていないと、ひらがなで始まる語、ひらがな混じりの語は抽出されない。

③1文字辞書

1文字でもキーワードとして抽出したい語を登録する。ここに登録されていない1文字の語は抽出されない。人名、地名などの抽出に使用する。

④同義語辞書

表記上の違いなどで検索結果が異なることを避けるため、全く同じ意味を持つキーワードを対で登録しておく。

なお、(1)から(3)までの辞書が、文書からキーワードを抽出する際に参照されているのに対し、同義語辞書は、検索する際に使用する辞書なので、データベースの容量が増えることはない。

(2) 書誌情報として入力するデータ項目

キーワードがデータ本文から切り出される場合、本文にないワードはキーワードにならないため、検索に不便な場合があると考えられる。たとえば、記事文中では“本県”とか“わが国”とかの記述がみられ、“日本”とか“県名”が省略されている場合が多く、後に“県名”などをキーとして検索しても、そのデータは出てこないことになる。そこで、それら不足していると考えられる項目をなんらかの形で補ってあれば、ユーザーにとってより便利となる。

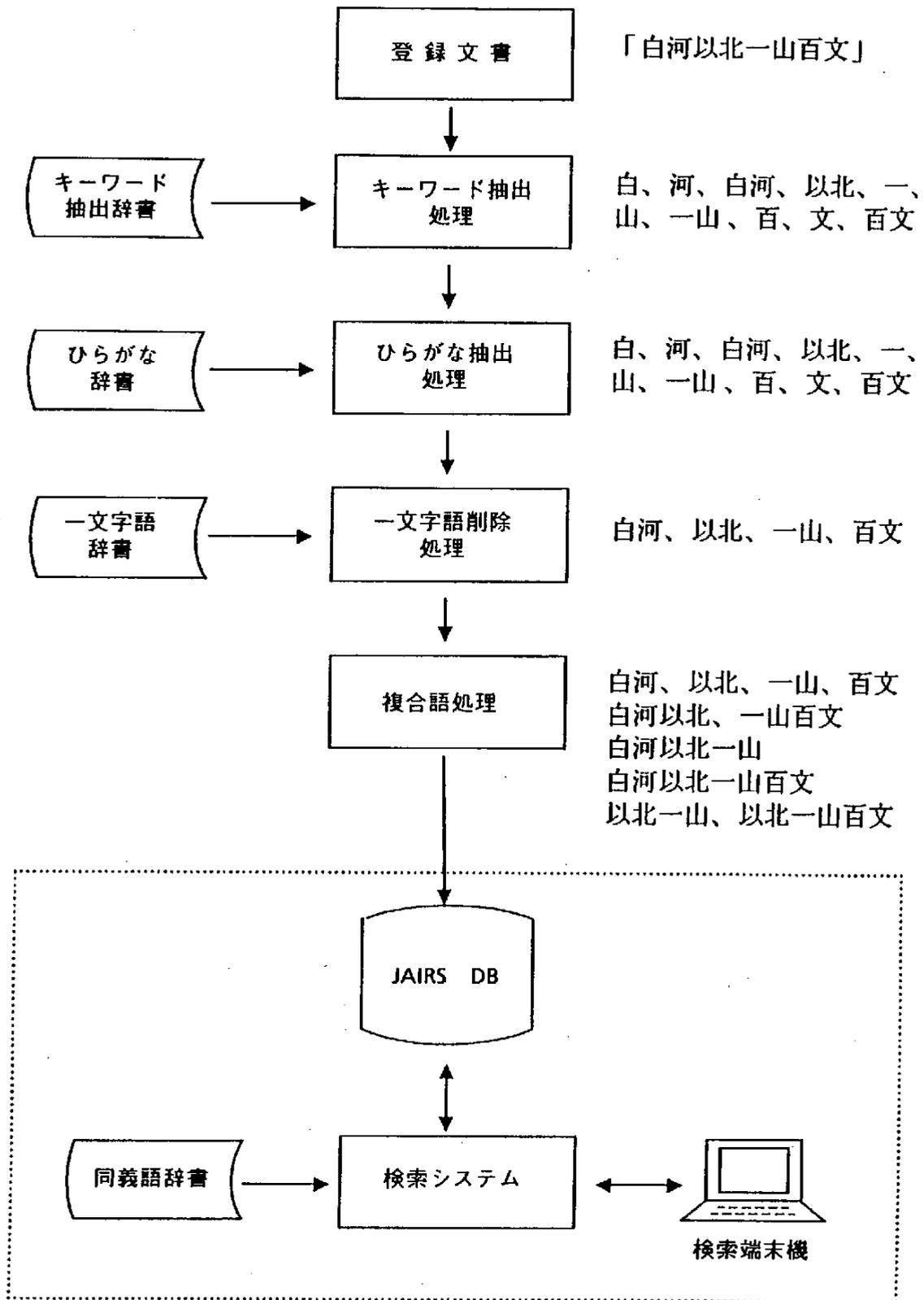


図 4-5 キーワード抽出処理の流れ

受託社は、その項目をデータ項目と称して入力し、データに付加している。そこで考えられた項目は、掲載日、発行元、発信元、面名、版別、記事の分野(分類)、地紋見出し、著作権問題のクリアなど公開の可否、国・県名、訂正記事の有無、写真の有無などである。

そのパソコン画面上のイメージを図4-6に示す。図4-6のデータ項目のうち、分類コード、見出し、追加キーワード、国別コードは、オペレーターがパソコンのキーボードを操作して入力する。他の項目は、システムが自動的に入力する仕組みになっているので、オペレーターは、たとえば“自社ものだが著作権がない”など例外だけを処理する。

4.2.3 データベース構築作業

ここで実際のデータベース構築作業の概略を述べる。

この報告書作成の時点では、部門外へのデータベースサービスを行っておらず、午前10時から端末4台を用い、午後4時ころ1日分の朝・夕刊の蓄積作業を終了している。

(1) 切り抜き作業とデータ蓄積

①切り抜き作業

まず、パソコン画面上に、紙面イメージ呼び出す。(図4-7)。次いで画面を見ながら、マウスを操作し、カーソルを該当記事に当ててクリックする。これにより、1件の記事の切り抜きが終わり、その分の色が変わる。この操作を繰り返すことにより1カ面の切り抜き処理を完了する。

②書誌情報入力

モードを変更し、画面上に記事(図4-8)を呼び出して、先のデータ項目画面(図4-6)を見ながら必要な情報を入力する。

③データ蓄積

切り抜きを終えれば、ただちにホストへ送ってデータ蓄積ができるが、現状では、1日分の作業を全て終えた段階で一括してデータを蓄積している。

(2) その他の作業

その他、関連した作業として、マシンのオペレーション、辞書のメンテナンス(特に同義語辞書のメンテナンス)作業、開発作業(検索システム、FAXサーバーの開発)などがあるが、ここでは省略する。

| データ項目 | | | | | |
|---------|--------------------------|--------|------|---|------|
| 記事番号 | K19920326A15X070 | | | | |
| 掲載日 | 19920326:平成4年3月26日 | | | | |
| 発行元 | K:河北新報 | | | | |
| 発信元 | 1:共同通信 | | | | |
| 面コード | A1:朝一 | 版別 | 5X | | |
| イメージ | 1:有 | | | | |
| 分類コード1 | 285:道路 | 分類コード2 | | | |
| 見出し | | | | | |
| 追加キーワード | <input type="checkbox"/> | | | | |
| 文字数 | 0967 | 訂正 | | | |
| 著作権 | 0:無 | | 0:無 | | |
| 非公開 | 0:無 | | | | |
| 国別コード | J10:東北 | | | | |
| 写真 | 00:無 | 図 | 01:有 | 表 | 01:有 |

◎

◎

◎

国道に102路線6061キロ昇格/道路審答申/東北は新たに6路線

道路審議会(高橋国一郎会長)は25日、山崎建設相に(1)主に都道府県道から102路線、延長6061キロを選び、新たに国道に指定する(2)一般国道の自動車専用道路のうち複数の国道からなる9路線に新番号を付ける一などを求める答申をした。

【地図】

【おことわり】「一般国道の追加指定路線(東北)」の表は記事データには入力されておられません。イメージ・データでご覧ください

図 4-6 データ項目

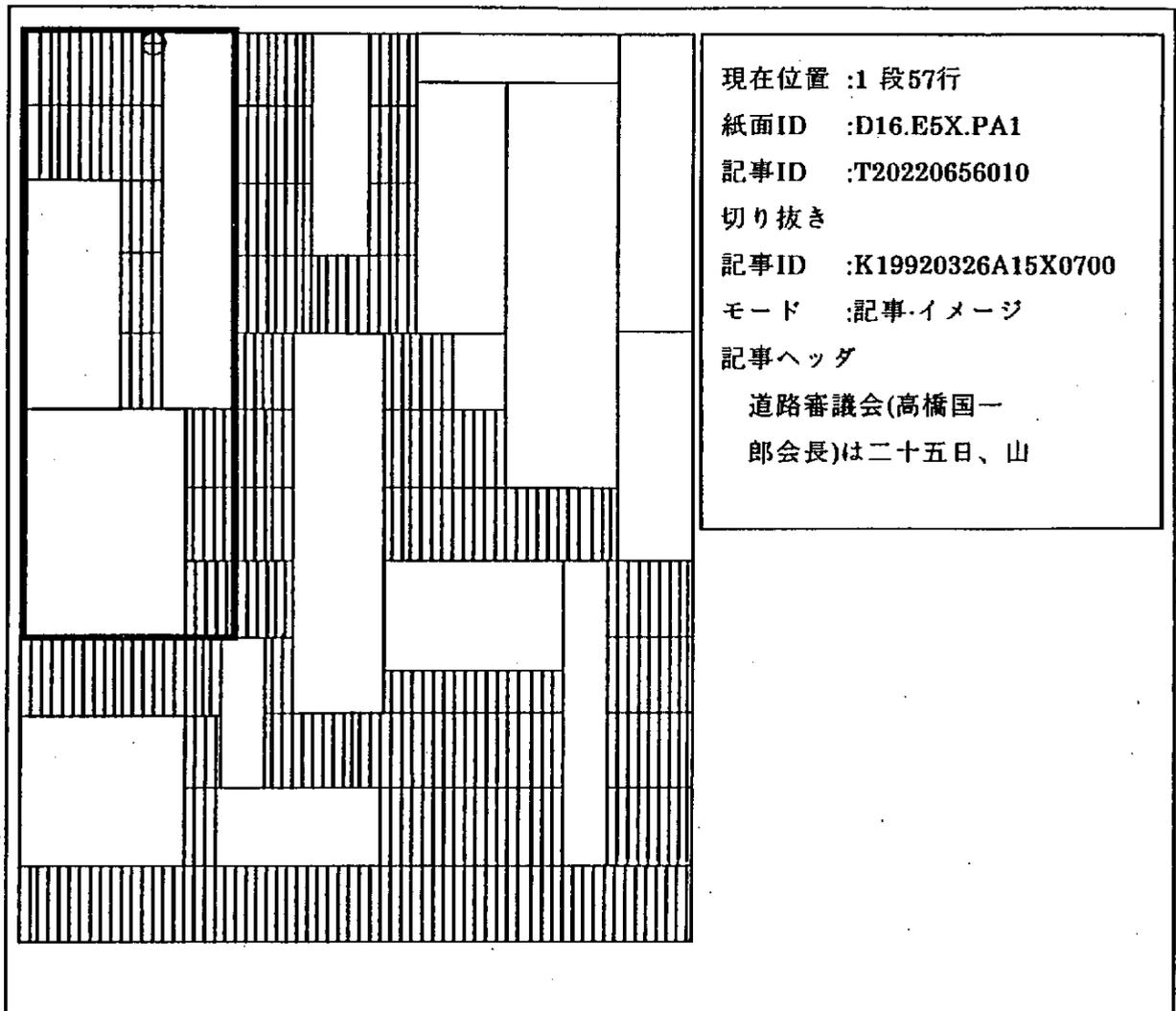


図 4-7 切り抜き作業画面イメージ

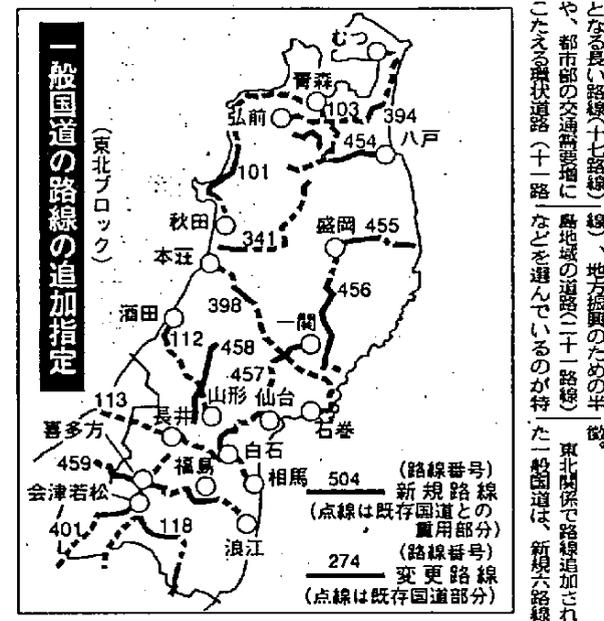
K19920326A15X070

国道に102路線6061キロ昇格

道路審答申 東北は新たに6路線

道路審議会(高橋国一郎会長)は二十五日、山崎建設(株)の提案に基づき、国道の追加指定を審議した。結果、新たに六路線を指定し、延長は六千六百一十一キロとなった。また、現行の国道の延長は、延長千口(延長)はこれまで最大規模。三十一日に閣議決定し、知事が道路の詳しいルートを決めて来年四月から新国道がスタートする。

今回の追加指定で国道の総延長は建設省が定めた目標の五万キロを超え、五万三千四百四十九キロとなる。



| 路線 | 起点 | 終点 | 追加指定区間 | 延長(キロ) |
|------|-------|--------|---------------------------------|--------|
| 101号 | 青森市 | 秋田市 | 八戸町～昭和町(秋田県) | 62 |
| 103号 | 青森市 | 大館市 | 青森市～十和田湖町(青森県) | 36 |
| 112号 | 山形市 | 酒田市 | 鶴岡市～酒田市 | 33 |
| 118号 | 水戸市 | 会津若松市 | 須賀川市～下郷町(福島県) | 51 |
| 341号 | 鹿角市 | 本荘市 | 協和町～本荘市(秋田県) | 41 |
| 394号 | むつ市 | 弘前市 | 六ヶ所村～七戸町(青森県) | 30 |
| 398号 | 石巻市 | 本荘市 | 湯沢市～東由利町(秋田県) | 28 |
| 401号 | 会津若松市 | 沼田市 | 会津若松市～南郷村(福島県) | 59 |
| 454号 | 八戸市 | 青森県大鰐町 | 八戸市～十和田湖町(青森県) 秋田県小坂町～青森県大鰐町 | 84 |
| 455号 | 盛岡市 | 岩手県泉町 | 盛岡市～泉町(岩手県) | 93 |
| 456号 | 盛岡市 | 宮城県本吉町 | 岩手県紫波町～宮城県本吉町 | 107 |
| 457号 | 一関市 | 白石市 | 一関市～白石市 | 160 |
| 458号 | 新庄市 | 上山市 | 新庄市～上山市 | 89 |
| 459号 | 新潟市 | 福島県浪江町 | 新潟市津川町～福島県浪江町 | 159 |

と変更(延長八路線たうた)地図、等。

新規では、一関市から仙台市の西部を経て白石市に至る国道101号や、盛岡市から八戸市を通過して宮城までそれぞれ延長された。八戸市に至る454号など追加指定された八路線の延長は、六路線合計追加指定区間の総延長は三百四十キロ。

今回の追加指定では新規、変更とも既存国道と重複する部分は除いてある。国道は六五年の道路法改正で、二百二十二路線、二万七千五百五キロが定められ、六九年から八一年まで四回追加指定されている。(20面に関連記事)

図 4-8 切り抜き記事呼出イメージ

4.2.4 システムの特徴

(1) 検索システム

検索システムは、ほぼ他社と同じ方式だが“漢字”・“ひらがな”による検索方式を採用、3月外部仕様書の作成、7月開発終了、8月サービス開始という経過をとっている。

(2) 受託社のデータベースシステムの特徴

以上述べてきた受託社のデータベースシステムは、およそ次のような特徴を持っているといえることができる。

①新聞製作システムからのデータの自動取り込み

②電子的切り抜き方式の採用

③紙面のデータベース化の範囲の広さ

掲載された記事のほとんど全てをデータベースとして収録している。

④基本的にはキーワードを自動的に抽出

記事文中から自動的にキーワードを抽出することができる。

⑤自由語検索（フリーターム）が可能

思いついたままの言葉（記事文中のキーワードはほとんど全て指定可能）で検索ができる。

⑥検索キーワードに漢字の利用が可能

キーワードとして漢字の指定ができる。

⑦豊富なデータ項目に基づく検索が可能

分類、国別コード、表の有無、図や写真の枚数、追加キーワード等の豊富なデータ項目によって、高度な検索が可能である。

⑧同義語辞書の装備

同義語辞書を装備しており、関連するキーワードでの記事検索が容易になる。例えば、ニューメディアという言葉について、パソコン通信、CATV、キャプテンといったような辞書を定義しておけば、キーワードは“ニューメディア”のみを指定することにより定義された4つのキーワードについての検索が可能となる。

⑨切り抜きイメージをそのままFAXに出力

場所を問わずに、新聞の切り抜きをファクシミリで提供できる。

⑩G4FAX対応品質でイメージデータを光ファイルに蓄積

鮮明な画質でのイメージデータの提供が可能である。

以上、受託社のデータベースシステムの概要を述べたが、ここまでの説明では、外部からデータベースを利用する仕組みにふれていない。

次章では、ここで述べた記事データベースへアクセスし、市中から利用するためのシステムについて述べる。

5. パイロットシステムの構築

5.1 記事データベースと「コミネット仙台」との接続の課題

5.1.1 配信システム構築の構想

4章で述べた記事データベースシステムは、主として社内でデータベースを利用するための仕組みであり、一般のユーザーが受託社の社外から利用できるようにするためには、まずデータを外へ出す仕組み（配信システム）を作る必要がある。具体的には、通信制御装置および必要なソフトウェアを導入し、アプリケーションソフトを開発して、電話などの回線を通じて外へ出す仕組みを作り、その仕組みを通じて、受託社のシステムとユーザーのシステムをつないでオンラインで利用する方法や、受託社からディストリビューターへ記事電文を送り、そのディストリビューターを通じてユーザーが利用するといった方法がよく知られている。

5.1.2 「コミネット仙台」のシステム概要

「コミネット仙台」のシステムは、(株)富士通の A-60 と呼ばれるコンピューターを中心に、MSIA (Multiple Serial Interface Adapter) 6 セットのほか増設メモリ、ディスクなどで構成されている。(図5-1参照)

なお、1セットの MSIA は、4 回線を制御するので、パソコンユーザーとの間の回線は、24 回線となる。

5.1.3 接続のための課題

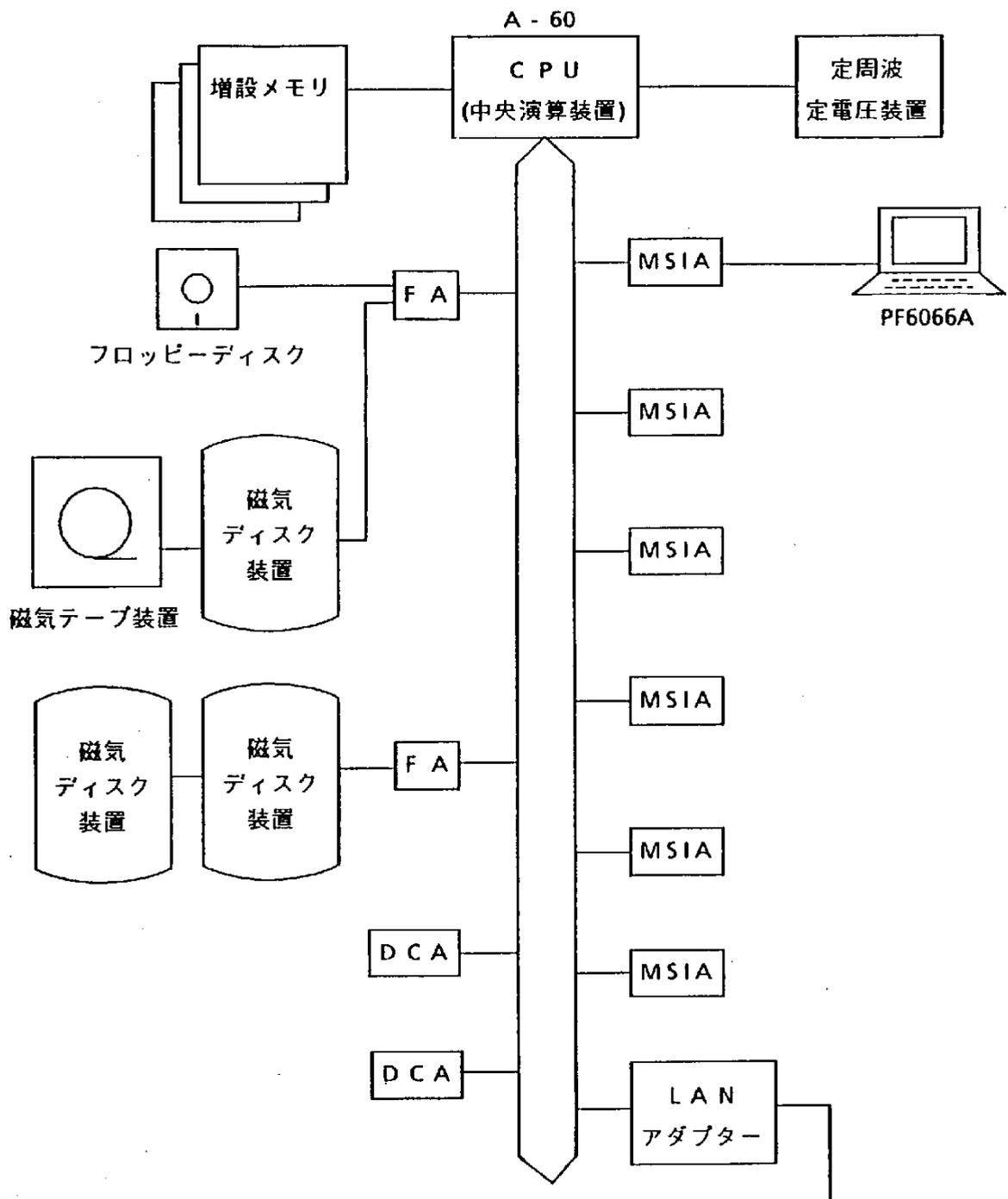
上記のような、A-60 を中心としたシステムと、IBM 社のコンピューター 9121/320 を中心としたシステムを接続した例はまだない。

先行しているシステムとして、北海道新聞社のシステムがある。同社は、先に IBM のコンピューター 3090 をホストとする CTS を実現し、その後 4381/R91 を中心としたシステムでデータベースを構築、平成3年秋、A-60 をホストとするパソコンネットワーク（オーロラネット）を通じてデータベースサービスを実用化した。(受託社調べ)

しかしながら、CTS およびデータベースのシステム構成が異なっているほか、ソフトウェアに自社開発部分が多く、システムをそのまま受託社へ置き換えることは簡単でないと思われる。

A-60 と IBM9121 の両システムを接続しようとした場合、当面解決を要する技術的課題として、

- ①両システム間の通信方式、通信手順の決定
- ②通信方式・手順に基づいてソフトウェアを開発するためのツールの開発



MSIA : マルチ・シリアル・インターフェイス・アダプター
 FA : ファイル・アダプター
 DCA : データ・コミュニケーション・アダプター

図 5-1 「コミネット仙台」のシステム構成図

- ③通信制御装置のハードおよび基本ソフトの導入
- ④アプリケーションソフトの開発（文字のコードやフォーマット変換含む）
といったことが考えられる。

5.2 パイロットシステム

前述のように、この報告の時点では、受託社は記事の蓄積作業を続ける一方、検索システムの開発を着手した段階であり、上記のような接続の基本的な課題を加えて、それらを短期間に一挙に解決することは困難である。そこで、まず通信上の主要な課題を解決して、パイロットとなるシステムを構築、そののち早期に本システムを構築することとした。

構築したパイロットシステムの概要は次の通りである。

5.2.1 通信方式

両社間の通信方式を、CCITTのX.25とし、第1種パケット交換回線（DDX-P、9,600 bps、パケット多重化機能8チャンネル）1回線を開設した。

5.2.2 通信制御装置の概要

(1) 名称 IBM3745/170

(2) 概要

1985年5月に発表されたIBM社の通信制御装置で、チャンネル・アダプター4つを持ち、最大112回線と接続できる。これを、1986年発表された前身の3720と比較した場合、条件によっては、約4倍の回線数とトランザクション量をサポートできるとされている。

実際に設置したのは、最大伝送速度19.2 Kbpsで16回線に対応可能な構成になっている。

(3) 特徴

特徴として、以下のような点が挙げられる。(受託社調べ)

- ①IBMホスト・プロセッサにチャンネル接続（3 MB/s 転送）
- ②最大112回線（19 Kbps以下）の制御が可能
- ③高速デジタル回線（ISDN）に対応（1.5 Mbps）
- ④トークンリングLANに対応（最大2）
- ⑤各種プロトコルに対応（X.25、ISDN、SNAなど）
- ⑥自動パワーオン/リスタート機能（遠隔地で有効）

5.2.3 システムの接続とネットワーク・ソフト

(1) 両システムの接続は、受託社のコンピューター IBM9121/320、通信制御装置 3745/170、DSU (Digital Service Unit)、NTT の回線 DDX-P、DSU、A-60 といった構成になっている。(図5-2)

(2) 基本ソフトの導入

端末、プログラム間の通信を制御する機能、通信ネットワークの制御といった基本ソフトを導入したほか、X.25 を IBM の世界にフィットさせるためのツールを開発、導入した。(技術編6. アプリケーションシスラト構成参照)

(3) アプリケーションソフトの役割

アプリケーションとして開発するプログラムの主な役割はおよそ以下の通りであり、本システム完成まで開発、テストが続けられ、それがネットワーク上でも確認される必要がある。

- ① IBM 漢字コードとシフト JIS コードの変換を行う
- ② コマンドの解析を行う
- ③ 要求されたコマンドの処理を行う(検索、見出し表示、記事表示、FAX 出力など)
- ④ FAX 出力処理の制御を行う
- ⑤ エラー時の対応を行う
- ⑥ 会計処理(課金情報)を行う

5.2.4 パイロットシステムのテストと確認

以上述べたパイロットシステムは、両システムを接続して「プロトコル接続確認テスト」を終了することによって、当面の役割を終え、引き続いて本システム構築に入り、検索システムなどの進捗に伴ってテストが続いている。

「プロトコル接続確認テスト」を目的として、数次にわたって実施、確認したテスト項目は以下(表5-1)の通りである。

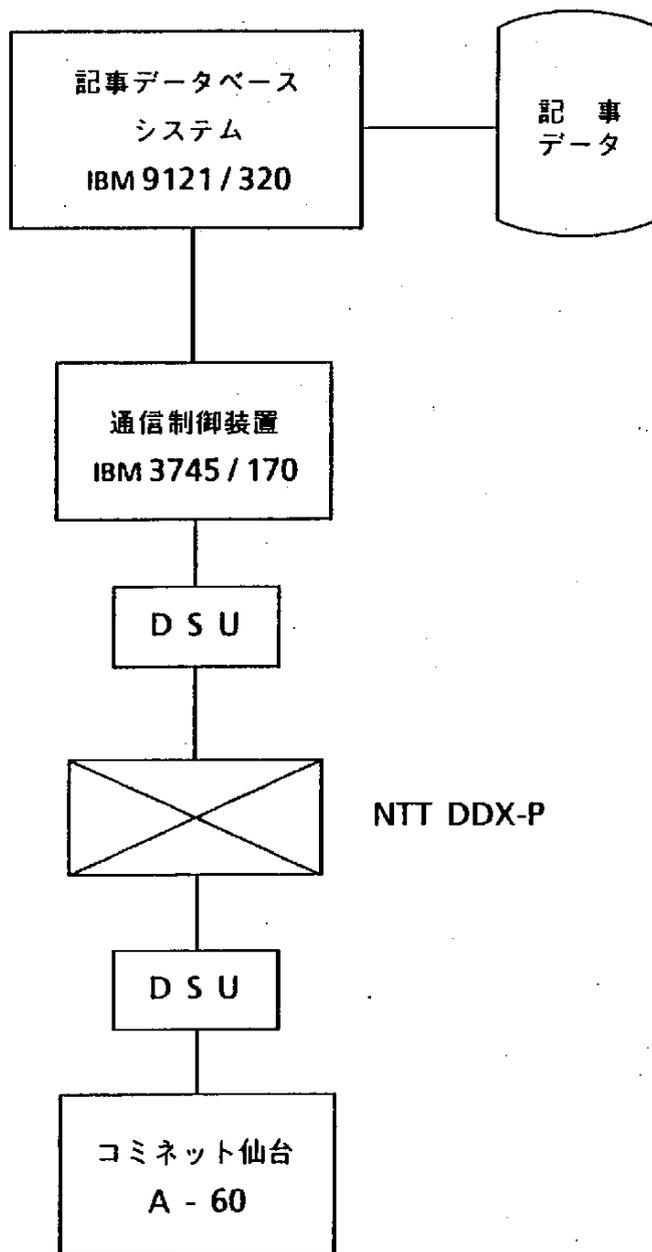


図 5-2 記事データベースシステムと「コミネット仙台」との接続

表 5-1 「コミネット仙台」との接続に伴うテスト確認一覧

(1) 正常ケース

| 番号 | 項目 | 検証方法 | コメント |
|----|------------------|---|-----------------|
| 1 | 800バイトを10回に分け送信 | CICS側のSEND時のRESPを確認する。 S/S端末に正しく表示できるかを確認する。 | パフォーマンス比較 |
| 2 | 10Kバイトを128回に分け送信 | CICS側のSEND時のRESPを確認する。 S/S端末に正しく表示できるかを確認する。 | パフォーマンス比較 |
| 3 | 32Kバイトを410回に分け送信 | CICS側のSEND時のRESPを確認する。 S/S端末に正しく表示できるかを確認する。 | パフォーマンス比較 |
| 4 | 800バイトを1度に送信 | S/S端末に正しく表示できるかを確認する。 | パフォーマンス比較 |
| 5 | 10Kバイトを1度に送信 | S/S端末に正しく表示できるかを確認する。 | 送信時の最大値を 求める |
| 6 | 32Kバイトを1度に送信 | S/S端末に正しく表示できるかを確認する。 | 送信時の最大値を 求める |
| 7 | 複数回に分けて受信 | CICSの1回目、2回目のRECEIVEのRESPを確認する。 | |
| 8 | 2台で同時に使用 | S/S端末が2台とも正しく作動するかを確認する。 | |

(2) 例外ケース

| 番号 | 項目 | 検証方法 | コメント |
|----|-------------------------------|---|------|
| 1 | S/S端末でデータ入力中の 回線切断 | CICS側のRESPを確認する。 | |
| 2 | S/S端末でデータ入力中の 電源遮断 | CICS側のRECEIVEのRESPを 確認する。 | |
| 3 | S/S端末でBREAKキー入力 | S/S端末の状態を確認する。 CICS側のRECEIVE時のデータ 及びRESPの状態を確認する。 | |
| 4 | S/S端末でデータの表示中に BREAKキー入力 | CICS側のRECEIVEのRESP 及びS/S端末の状態を確認 する。 | |
| 5 | ENTER押下後に BREAKキー入力 | CICS側のRECEIVEのRESP 及びS/S端末の状態を確認 する。 | |
| 6 | S/S端末でデータの表示中に CTRLキー+Cを入力 | CICS側のRECEIVEのRESP 及びS/S端末の状態を確認 する。 | |
| 7 | S/S端末でデータの表示中に CTRLキー+Sを入力 | CICS側のRECEIVEのRESP 及びS/S端末の状態を確認 する。 | |
| 8 | CICSで送信権なし時にSEND | 2回目のSENDのRESPをを確認 する。 | |
| 9 | CICSで送信権あり時に RECEIVE | CICS側のRECEIVEのRESP を確認する。 | |

| 番号 | 項目 | 検証方法 | コメント |
|----|----------------|----------------|---------------|
| 10 | CICSで送信権あり時に終了 | S/S端末の状態を確認する。 | |
| 11 | CICSで送信権なし時に終了 | S/S端末の状態を確認する。 | リカバリの手段を確認する。 |

6. 今後の展望

6.1 本システムの実現へ向けて

パイロットシステムの目的は、本システムの実現を容易にすることにあり、早急に本システムを構築し、サービスを開始していくことが望まれる。

パイロットシステム構築後、本システムの実現へ向けて解決が必要とされる課題は、およそ以下の通りである。

(1) 検索に対してのコマンド体系統一の必要性

受託社側が「コミネット仙台」のコマンド体系に準拠するにあたって、数件の対応不可能なコマンドがあり、対処が必要となる。

(2) 「コミネット仙台」のユーザーへの FAX サービスの提供

手動でサービスを行うことも可能であるが、FAX サーバーを使用しホスト側でアプリケーションを開発、社内検索システムと共用することが望ましい。

(3) FAX サービスに対する課金方式の決定

次のような複数の課金方式があるが、それぞれに課題が挙げられる。

- ① 送信した用紙枚数での課金（イメージの大きさにかかわらず、同じ1枚として課金される）
- ②ダイヤルQ2を使用した課金（開発のワークロードが大きいと予想される）
- ③通話時間での課金（距離による価格設定が必要となる）

(4) データベース利用の課金方式

次のような課金方式があり、検討を要する。

- ①送信した記事数で課金（記事の大きさに関係しない）
- ②検索時間で課金
- ③受託社側の資源使用率で課金

(5) 「コミネット仙台」のユーザーのレベル分けの課題

記事を公開できる範囲や FAX サービス利用の有無により、以下のようなレベル分けが考えられ、ユーザー ID の事前登録などで対処することが必要となる。

- ① FAX 利用不可の一般ユーザー
- ② FAX 利用可能な一般ユーザー
- ③受託社ユーザー（受託社の記者などで、FAX 利用を可能とし、検索レベルを

高く設定する)

- (6) 「コミネット仙台」のユーザーに対する“お知らせ”情報の提供
受託社側で情報を事前に登録し、プライマリーメニューの表示前にその情報を流すことができる。
- (7) 「コミネット仙台」のユーザーへのイメージ表示機能の提供
当面、困難と思われる。

6.2 システム発展の展望

当面の目標である本システムが実現し外部へ配信する仕組みができたとして、受託社のデータベースには

- ①アクセス可能なネットワークをどのように拡大していくか
- ②パソコンなど情報機器の操作が不得意なユーザーのニーズにどう応えていくかサービスの質および量をどのように充実させていくか

という基本的な課題がある。以下、これらの点について述べてみたい。

- (1) 記事、データベースの普及と多方面のユーザーニーズに応えていくには、アクセス可能なネットワークを拡大していく必要がある。そのためには、データベースシステムと新たなネットワークとの接続、配信先のゲートウェイ機能によるネットワークの拡大などが考えられる。

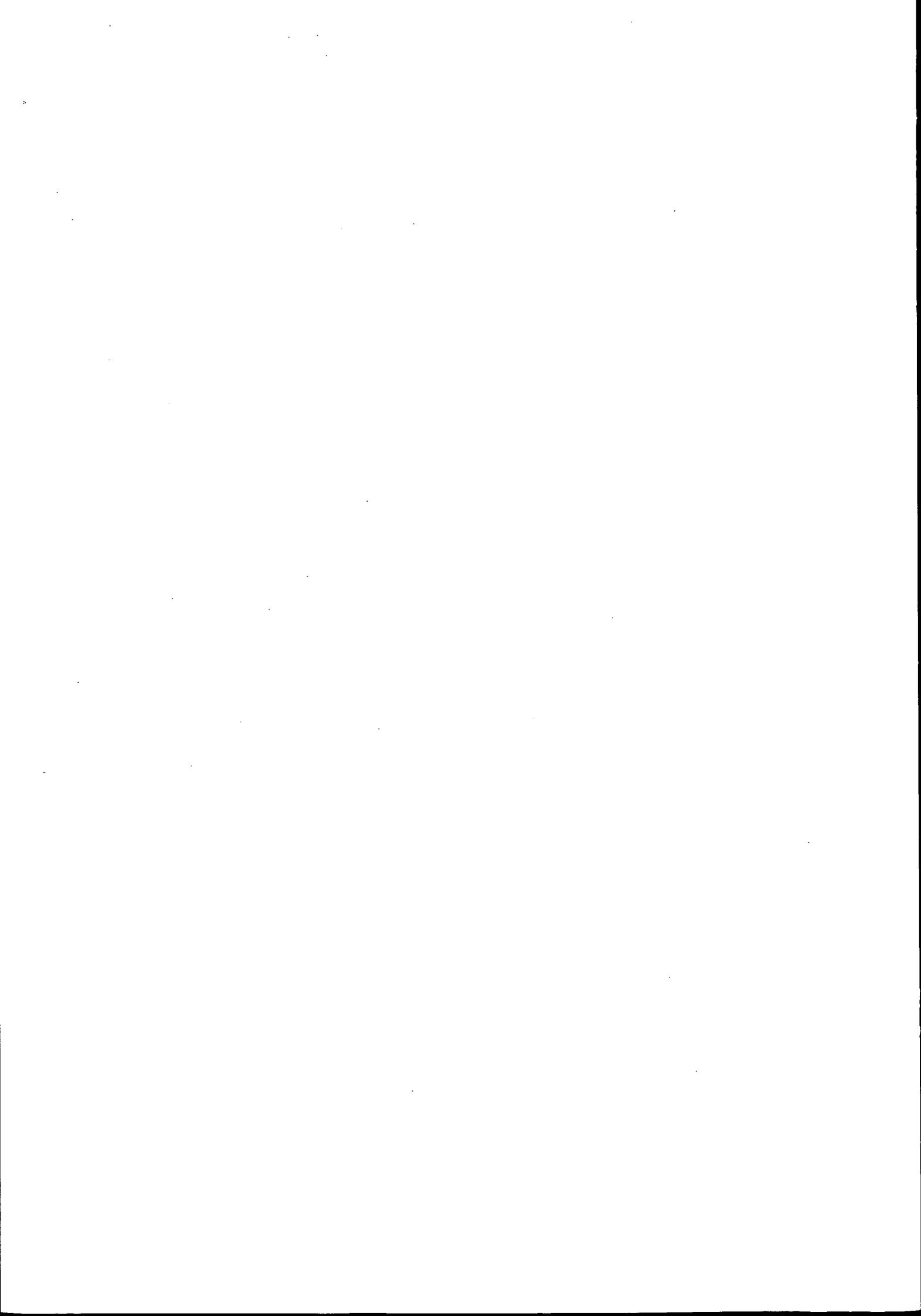
- (2) 「パソコン通信を使ってデータベースサービスを誰もが利用できるようになるには5～10年かかるだろう。なぜなら、そのころになればパソコンで教育を受けた世代が、社会人としてコンピューターの活用を始めるからである」という説がある。

しかし、私達はそれまで手をこまねいて見ているわけにはいかない。そこで、これまでキーボードに触れることのなかった世代や、触れることを拒んでいた世代を巻き込んでの啓蒙活動が必要となってくるだろう。

そのためには、マンマシンインターフェイスの向上を図ると同時に、啓蒙活動の第一歩として、きめの細かいデータベース教室や活用セミナーの実施、そしてイベントや雑誌媒体を用いてデータベース自体のピーアールを主体的に展開していく必要がある。ただ利用されないと嘆くのではなく、積極的に利用される環境を作り出していくことが必要と思われる。

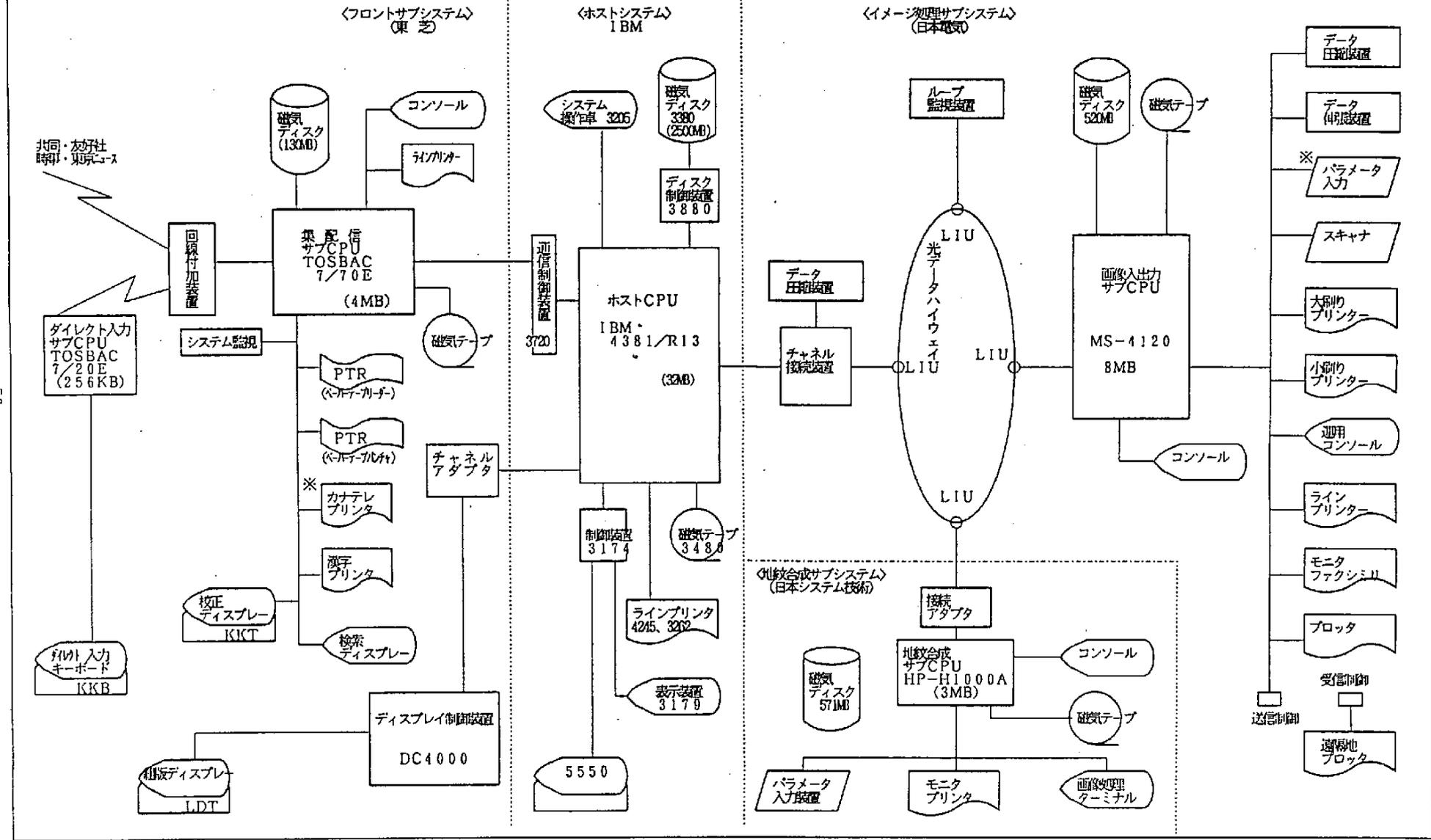
- (3) 現在、データベースの検索端末として、パソコンが一般的に利用されている。しかし、将来的には、例えば、CATVから発展していったテレビ端末機の活用など、操作性に優れた新しい情報機器の誕生が期待される。

- (4) データベースサービスの内容として、当面は掲載記事に限られた提供となるが、段階を踏まえ「聞き取り調査」に基づく要望事項について、実現を図っていく必要がある。
- (5) 新聞は分野別に記事が掲載され、読者は関心のある記事についてその情報を入手していく。一方、記事データベースは、様々な分野の記事が1つのまとまりとして存在し、ある情報を欲する人間がキーワードを基に最終的に入手したい情報を探り出していく。つまり、新聞と記事データベースサービスは、次元の異なった商品と言えよう。その意味で、記事データベースサービスは、既存の資源を活用した新しい商品の提供であり、既存の未利用資源を更に活用することにより、サービスの質及び量の一層の充実が図れるものと考えられる。
- (6) 将来において、データベースと呼ばれるものが、現在と同じ形態、同じ概念で存在するかどうかは、大きな疑問である。現在では、想像も出来ない全く異質なものが、データベースと呼ばれる可能性もある。しかし、データベースはデータが蓄積されてこそ大きな意味を持つ。その意味で今データベースへの取組を開始することによって、将来の形態の如何んを問わずその時のサービスの主体として存在する可能性は非常に高い。



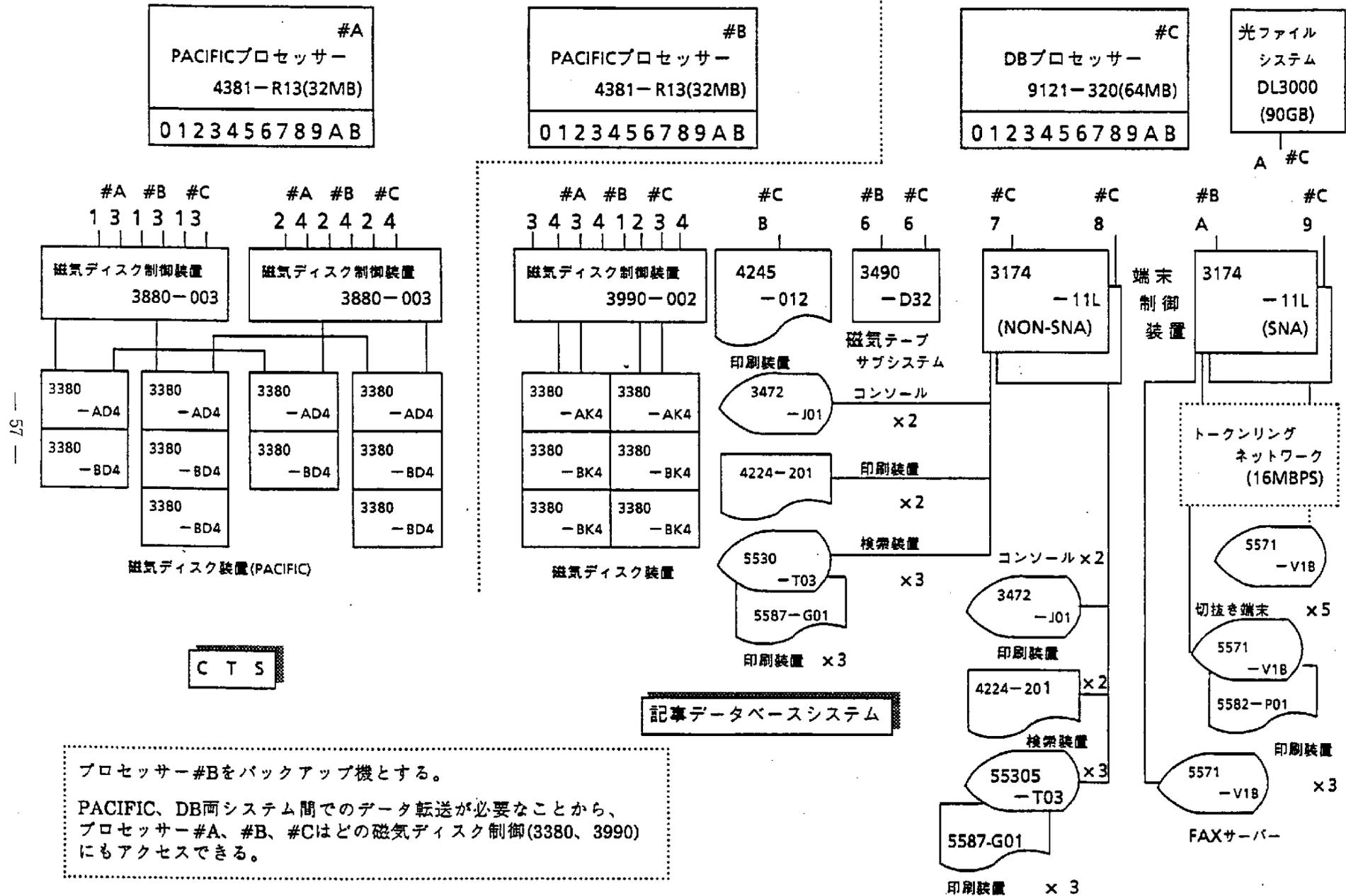
III 記事データベースアクセス用パイロット システムの構築に伴う技術資料

1. パシフィックシステム (CTS) 構成図



※ カナテレプリンタ: カナ文書用のテレタイプ
 パラメータ入力: 写真・図等の印刷情報の入力

2. CTSと記事データベースシステムの接続図 (次ページの機器概要参照)



◇ 記事データベースシステム機器概要

(1) 9121-320

中央演算処理装置

IBM が平成 3 年 9 月に発表した ES/9000 ファミリーでは、空冷式、単一プロセッサで最上位のモデル。記事データベースシステムでは、この機種をホストコンピュータとして導入し、入出力、端末装置などを制御する。性能は、現行 PACIFIC ホスト、IBM4381-R13 の約 5.5 倍であり、主記憶容量は 64 MB である。

(2) 3990-002

磁気ディスク制御装置

大容量の磁気ディスク装置を制御し、ホストプロセッサと磁気ディスク装置間のデータ転送の通路の役目を果たす。特に 3990-002 型は、4 つのデータ転送経路を提供する。

(3) 3880-AK4/BK4

磁気ディスク装置

大量の記事やイメージデータを蓄積する装置であり、端末機からこの中のデータを検索することになる。AK4/BK4 は 1 台で 7.5 GB あり、導入する 6 台では合計 45 GB の容量になる。

(4) 3940-D32

磁気テープサブシステム

制御装置とテープドライブを一体化したコンパクトな磁気テープ装置で、PACIFIC と完全互換性のある 1/2 インチカートリッジテープを使用する。

(5) 4245-012

印刷装置

インパクト方式の印刷装置であり、プログラムリスト、統計情報の印刷を行う。印刷速度は、最高 1,200 行/分である。

(6) 3174-11 L

端末制御装置

9121 プロセッサとその端末機とを、効率よく高速に接続する制御装置である。4 台設置するうち 2 台は、システムコンソール、検索端末機などを制御し、残りの 2 台でトークンリングネットワークを構築して、切り抜き端末機やレーザープリンターを接続する。

(7) 3299-002

端末多重装置

3174 制御装置の 1 本の同軸ケーブルを、効率よく端末機に分配する装置である。

(8) 8228-001

集線装置

ワークステーションなどの情報処理装置をスター状に接続するリング集線器であ

る。トークンリングネットワークに切り抜き端末機を接続している。

(9) 3472-J01 **Infowindow**

システム制御のためのオペレーター用コンソールである。

(10) 4224-201 **印刷装置**

インパクト方式の端末印刷装置である。印刷速度は、最高 200 文字/秒であり、コンソールプリンターとして使用する。

(11) 5571-V1B **PS/55**

32 ビット高速プロセッサ (25 MHz) を搭載した PS/55 シリーズの上位 OS/2 マシンである。作業性を重視し、大型 (20 インチ)、高解像度ディスプレイやマウスを接続している。6 台設置する中で 1 台は、切り抜きイメージをファクシミリ出力するためのサーバーとして使用する。

(12) 5582-P01 **印刷装置**

プリンター、スキャナー、プロッター、複写機の機能が一体になった印刷装置である。解像度は、400×400 ドット/インチであり、切り抜き作業の確認、検索結果の出力、イメージの入力など多用途に使用する。

(13) 5530-T03 **PS/55**

システム装置と 12 インチカラーディスプレイが一体となった端末機であり、社内における検索用端末として使用する。

(14) 5587-G01 **印刷装置**

半導体レーザーを光源とした電子写真方式のページプリンターである。解像度は、240×240 ドット/インチであり、データベースから検索した切り抜きイメージをプリントアウトする。

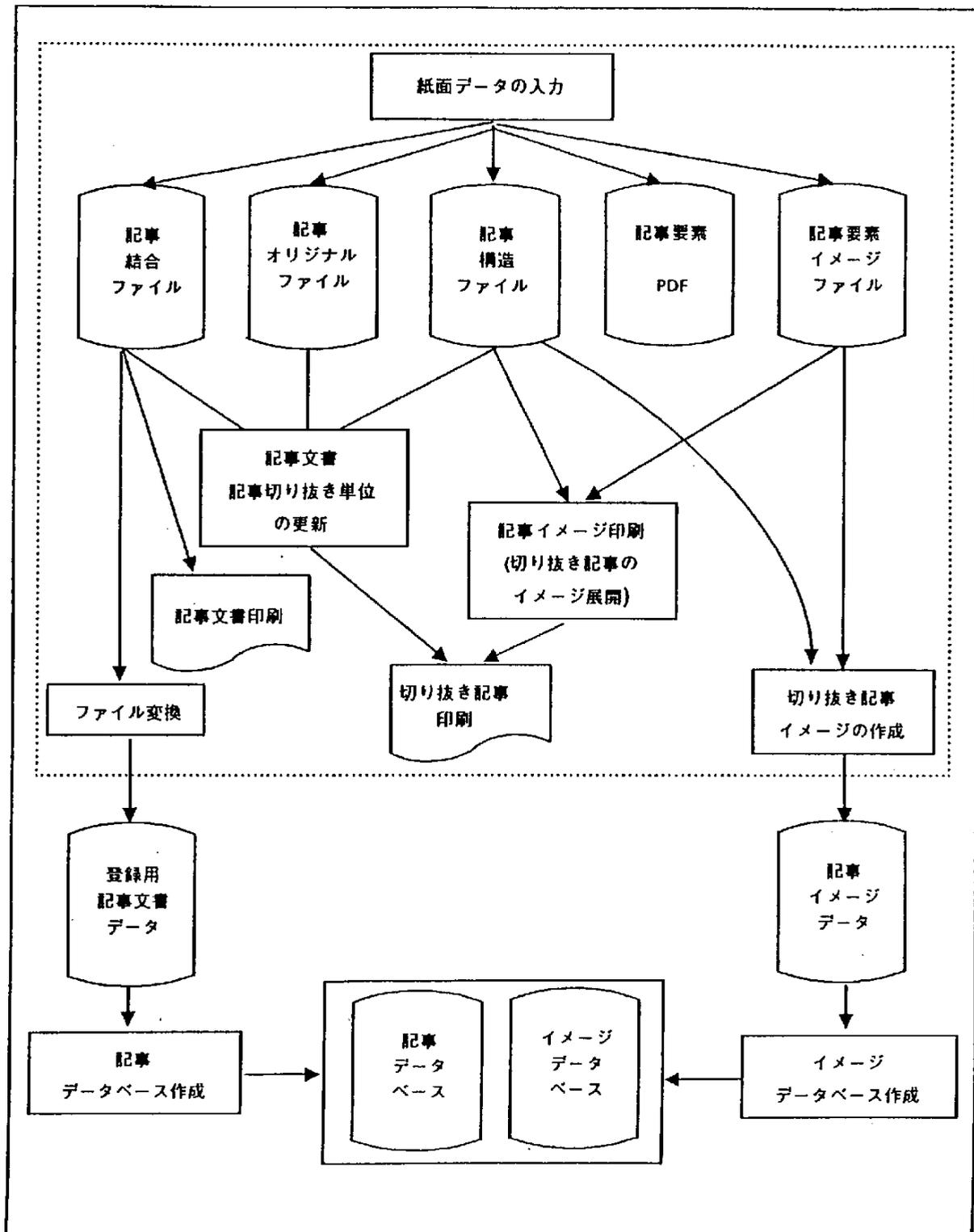
(15) DL 3000 **光ファイル装置**

兼松エレクトロニクスの光ファイルシステム。90 GB (650MB×150 枚)。記事データは、磁気ディスク装置に入ると同時に光ファイルシステムにも取り込まれる。

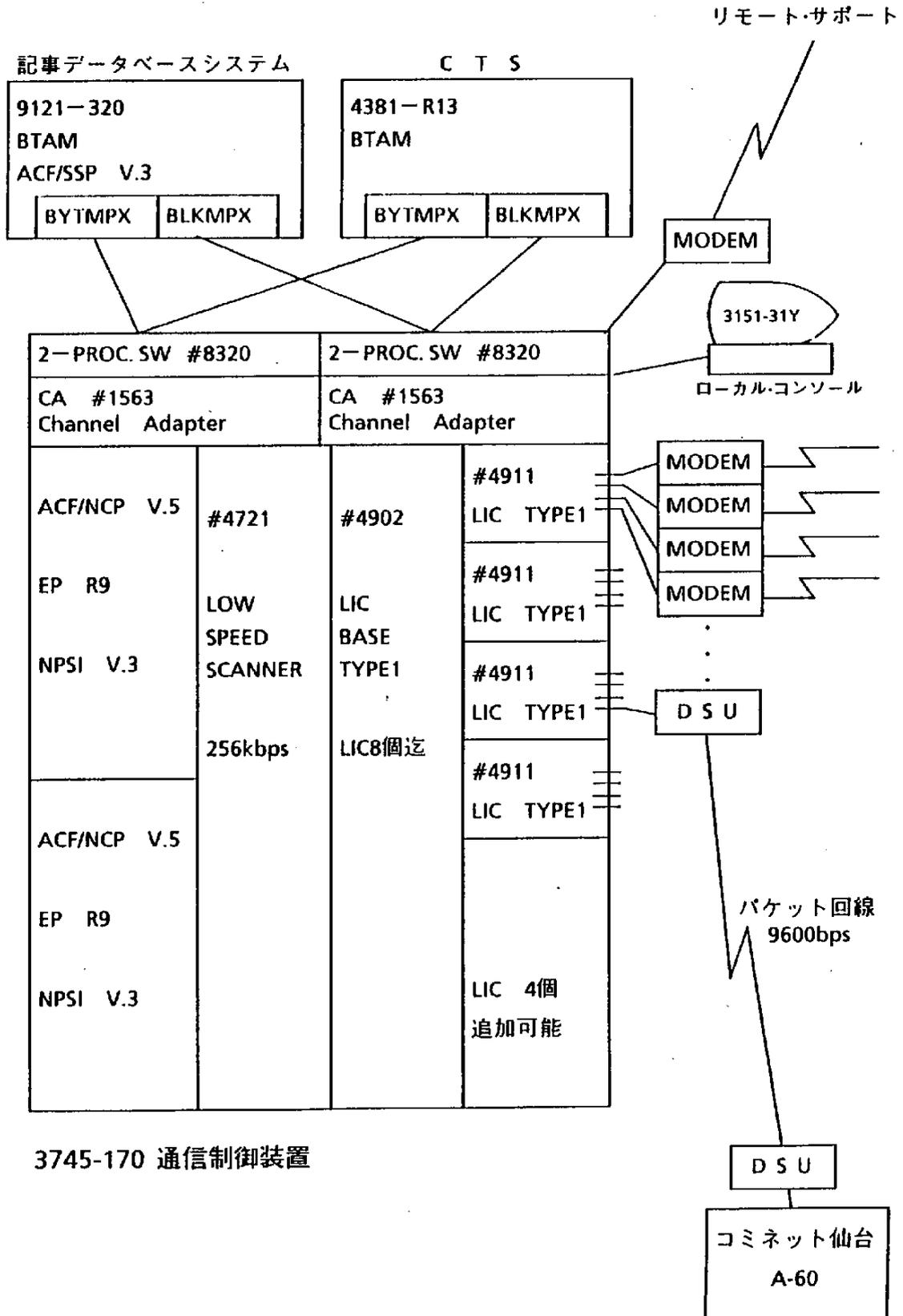
3. 記事データ蓄積処理フロー

CTSで作成された組版データを関連する記事単位に分割し、記事データベースとして蓄積していく。また同時に、紙面上に掲載された記事の形状を切り抜きイメージデータベースとして蓄積していく。

処理の主な流れは、次の通りである。



4. 通信制御装置構成図



◇ 通信制御装置機器概要

(1) 3745-170 は以下の接続が可能である

- ・ 4チャンネルアダプター
- ・ 最大112もしくは96中低速回線
112=96 (LIC 1, 3, 4)+16 (LIC 5)
96=64 (LIC 1, 3, 4)+32 (LIC 5)
- ・ 最大16のLIC 5 (32回線) 接続可能
- ・ 最大2高速デジタル回線 (TDM 経由: 1.5 Mbps まで)
- ・ 最大2トークリング

(2) LOW SPEED SCANNER (LSS: 低速スキャナー)

- ・ マイクロ・プロセッサによる最高256 Kbps までの回線走査機構
- ・ 最大6個の取付けが可能

(3) LIC BASE (LIB: 回線接続ベース機構) TYPE 1

- ・ LIC TYPE 1 の前提機構
- ・ 8個のLICの格納が可能
- ・ 最大3個の取付けが可能

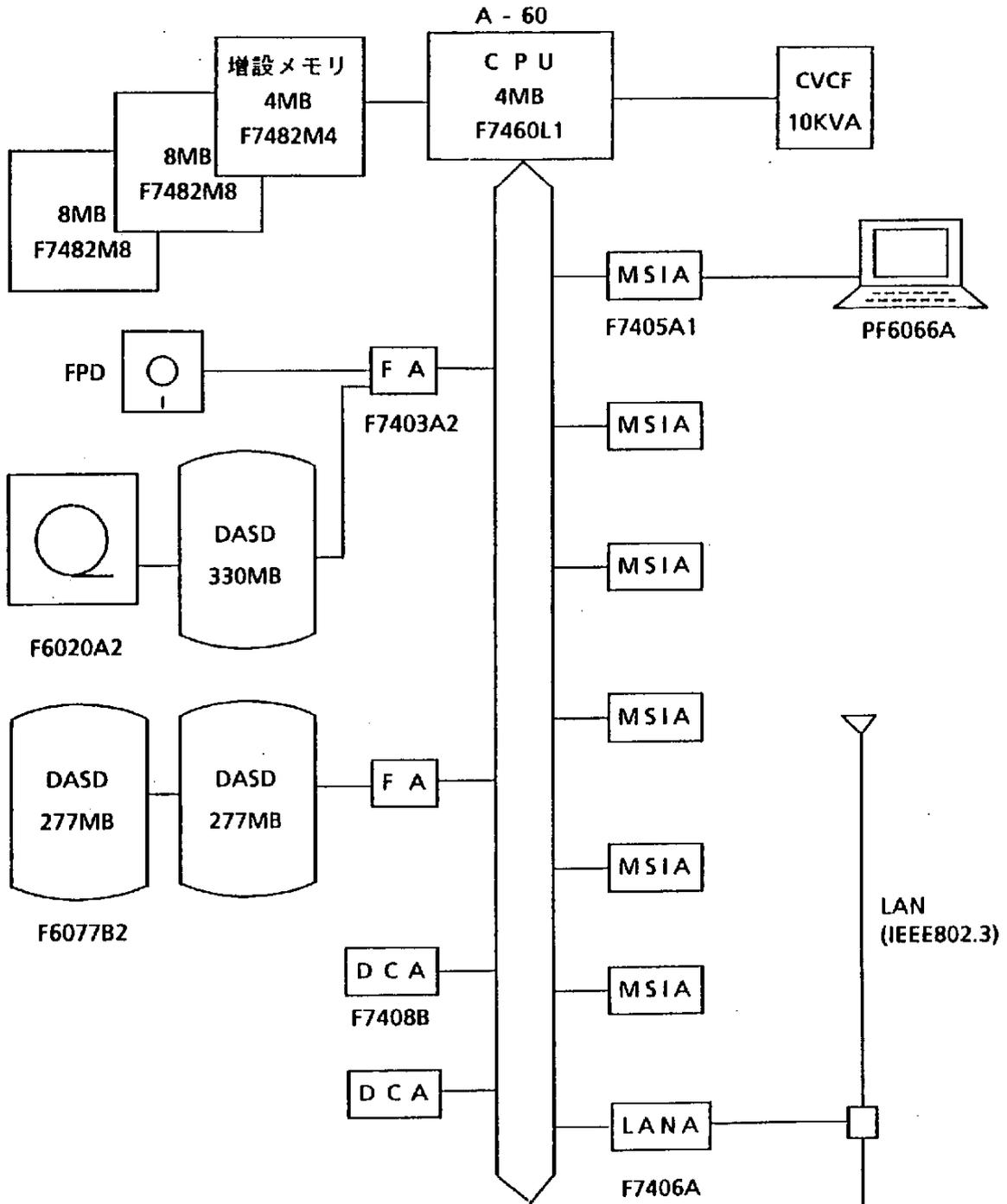
(4) LIC TYPE 1 (回線接続機構1型)

- ・ 4回線収容
- ・ ケーブルは4種類
NTT DCE 接続
NTT AUTOMATIC CALLING UNIT 接続
ASYN DTE 直接接続
SYN DTE 直接接続
- ・ 最大24個の導入が可能

◇ 通信制御装置 (IBM3745-170) の諸元

- 体積 : 750 mm(前) × 650 mm(横) × 1000 mm(高)
- 保守サービスエリア : 前 1000 mm
: 後 1000 mm
- 重量 : 230 kg
- 発熱量 : 0.75 KW
- 排気量 : 7.2 m³/分
- 消費電力 : 1.1 KVA
- 電源 : 単相 200~240 V
- 最大電流 : 6.2 A
- サイクル : 50 Hz または 60 Hz
- ストレージ : 4 MB
- ハードディスク : 72 MB
- キャッシュ : 32 KB
- 稼働環境
 - 温度 : 16~38°C
 - 相対湿度 : 8~80 %
 - 最高湿球温度 : 23°C

5. 「コミネット仙台」ハードウェア構成図

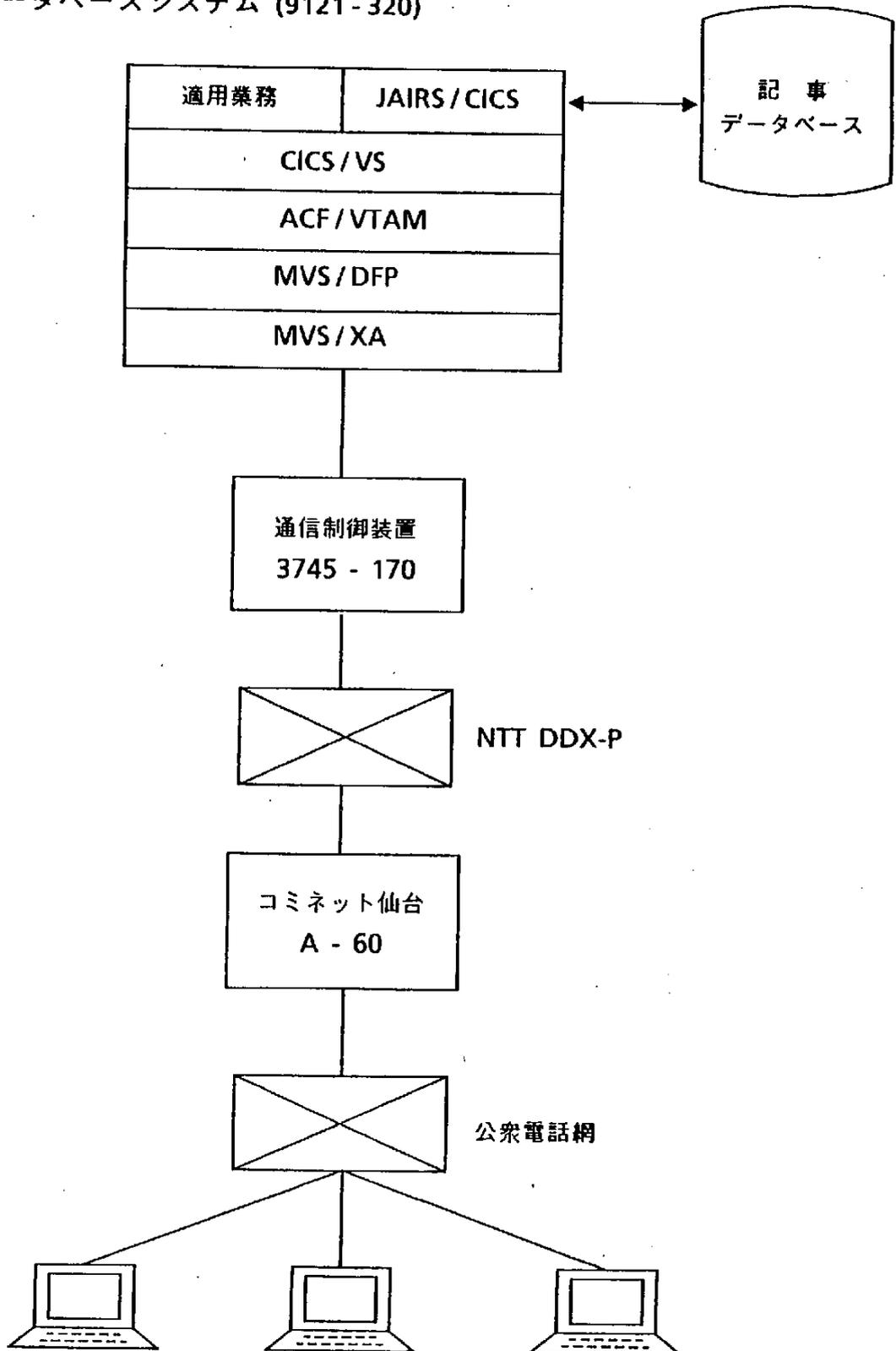


- MSIA :マルチ・シリアル・インターフェイス・アダプター
- FA :ファイル・アダプター
- DCA :データ・コミュニケーション・アダプター
- CVCF :定周波定電圧装置
- DASD :磁気ディスク装置
- LANA :LANアダプター

6. アプリケーションシステム構成

アプリケーションシステムの構成は、以下の通りである。

記事データベースシステム (9121-320)



◇ ソフトウェアと機能概要

(1) 基本ソフト

① MVS/XA (Multiple Virtual Storage/Extend Arctecture)

基本オペレーティングシステムであり、全てのソフトウェアの基盤となるものである。2Gバイトの多重仮想記憶領域を提供する。

② MVS/DFP (Multiple Virtual Storage/Data Facility Product)

MVS/XA の配下で、データ管理、装置管理、ライブラリー管理、ユーティリティー等の機能を提供する。

③ RMF (Resource Management Facility)

システム資源の使用状況、パフォーマンス情報を収集し、システム全体のパフォーマンス/キャパシティ管理に必要な情報を提供する。

④ DF/DSS (Data Facility/Data Set Services)

磁気ディスクのバックアップ及びリストア機能を提供する。

⑤ DF/SORT (Data Facility/Sort)

データの分類、組合せの機能を提供する。

⑥ TSO/E (Time Sharing Option/Extention)

適用業務開発者との対話処理機能を提供する。

⑦ ISPF/PDF (Interactive System Productivity Facility/Program Development Facility)

TSO/E の元で、メニュー形式により対話処理機能を提供する。

(2) 画像処理ソフト

① GDDM (Graphical Data Display Manager)

画像データ (新聞切り抜きイメージ) をコントロールする。
解像度変換、回転、縮小等の機能を提供する。

(3) データベース処理ソフト

① JAIRS/CICS (Japanese Information Retrieval System/Customer Information Control System)

記事データベースを制御するために、様々な機能を提供する。
データベースを制御する中核的ソフトである。

(4) プログラミング言語

① PL/ I (Program Language/one)

適用業務開発において使用する。アセンブラー言語と比較して高い生産性を提供している。

② ASSEMBLER (Assembler)

適用業務開発において使用する。PL/ I では対応できない場合に使用するものである。

(5) 提供業務プログラムの役割

- ① IBM 漢字コードとシフト JIS コードの変換を行う。
- ② コマンドの解析を行う。
- ③ 要求されたコマンドの処理を行う (検索、見出し表示、記事表示等)。
- ④ ファクシミリへの出力の制御を行う。
- ⑤ エラー時の処理を行う。
- ⑥ 会計処理 (課金情報処理) を行う。

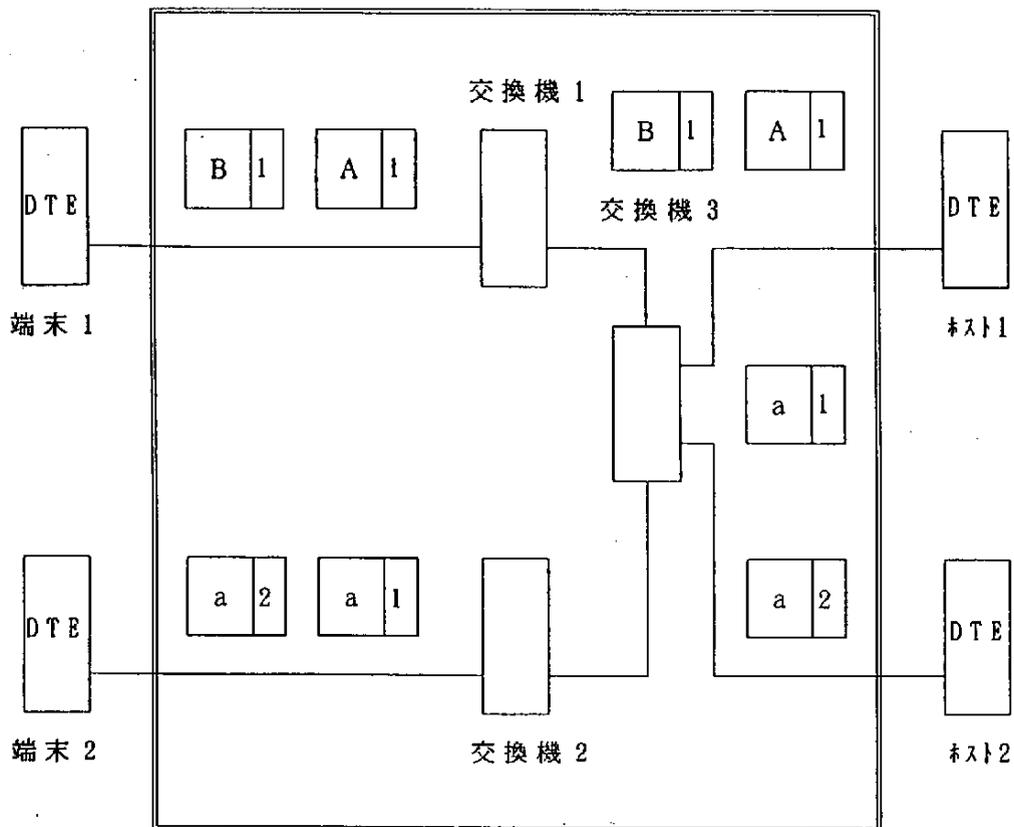
7. 通信仕様 (X.25 プログラミング)

7.1 X.25 パケット交換とは

7.1.1 概要

X.25 パケット交換とは、データ端末 (DTE : Date Terminal Equipment) からパケットと呼ばれる単位にデータを区切り、網内の交換機に蓄積され、そこでさらにあて先による交換機に中継線を経由して目的地の DTE 対してパケットを送信することにより通信を行う蓄積交換方式である。

この X.25 パケット交換は、国際勧告である CCITT X シリーズのいくつかの勧告で規定されている。



| |
|---|
| X |
|---|

 はパケットを表す。

| |
|--|
| |
|--|

 の部分がパケット・ヘッダーの部分であて先等の制御情報が入る。

| |
|---|
| X |
|---|

 の部分はユーザー・データが入る。

端末 1 から送り出されたパケットには、ホスト 1 へのあて先情報がセットされている。交換機 1 は、あて先情報を調べ、中継線の混み具合を見て交換機 2 か交換機 3 に対してパケットを送り出す。

7.1.2 X.25 パケット交換の特長

パケット交換は、従来の専用線や交換回線とは異なった次のような特長がある。

(1) 蓄積交換

パケット交換は、パケット交換機内の記憶域にいったんデータを蓄積し、あて先の情報が入ってるパケット・ヘッダーを調べて相手に送信する蓄積交換方式である。そのため、記憶装置上での何らかの処理を行うことが可能となる。

(a) パケット化

送信すべきメッセージを決められた長さに分割し、その各々に宛先などの制御情報を付与していくこと。

(b) パケット多重化

パケット伝送では、論理チャンネル上をパケット化したデータを送受信するが、1本の物理回線上に複数の論理チャンネルを持つことができる。

(c) 異速度通信

パケット交換では、蓄積交換であるためそれぞれの DTE と DCE (Data Circuit-terminating Equipment) が独立してパケット転送を行う。そのため、伝送スピードの異なる DTE 間の通信が可能である。

(d) 異手順端末通信

パケット交換機を仕様した蓄積交換であるため、記憶域内で DTE 間の通信手順の変換を行うことができる。

この機能を提供するのが PAD (Packet Assembly Deassembly) である。CCITT 勧告では、X.28 無手順端末 (スタート・ストップ) とパケット端末とが通信できるようなプロトコル変換を行う PAD を X.3 で規定している。

(e) 付加機能

NTT の DDX-P では、交換機内でさまざまな処理を行う付加サービスを行っている。

(f) 従量制課金

X.25 パケット交換では、交換機を通過したパケットの量に応じた課金を行っている。

(2) デジタル伝送

X.25 パケット交換では、デジタル伝送を行っている。

(a) 高速伝送

CCITT 勧告での最大速度は、64 Kbps となっている。

NTT の DDX-P では、48 Kbps が最大速度である。

(b) 信頼性

信号の減衰に対して再生中継器をおいて波形を整形しているため、雑音に対して強くなっている。

(c) 接続/切断時間

接続/切断時間は、1 秒以下とアナログ交換回線に比較して早くなっている。

(3) 接続形態

基本的には交換網であり、相手を自由に選択して通信する事が可能であるが、相手を固定的に接続する方法をとることもできる。

(a) 相手選択接続 (SVC : Switched Virtual Circuit)

CCITT 勧告では、VC (Virtual Call) となっている。

通信相手を自由に選択できる接続形態で、実際には呼の設定時に相手の DTE を選ぶ。つまり、発信を行う場合相手の回線番号をパケットに入れて DCE に対して呼設定を要求している。

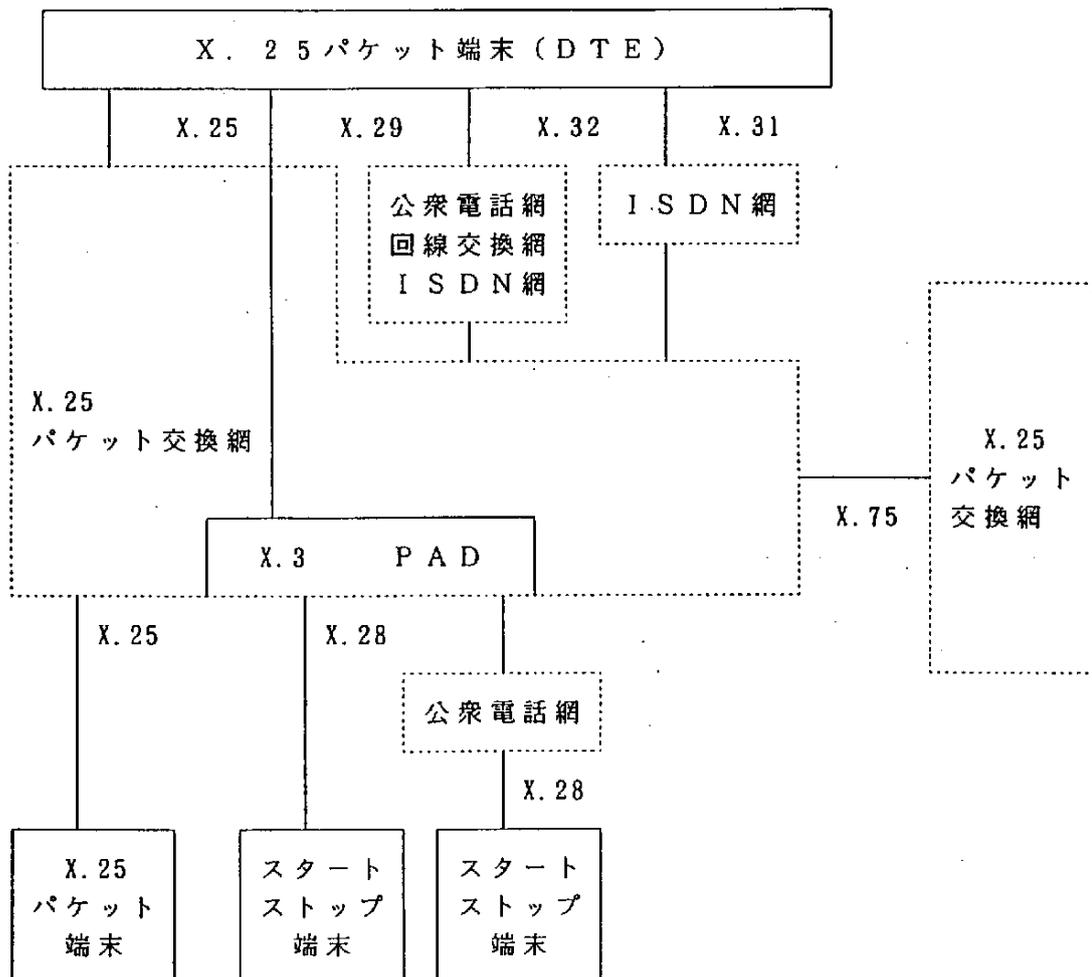
(b) 相手固定接続 (PVC : Permanent Virtual Circuit)

通信相手を網側で固定しておく接続形態で、通信相手は常に決まっている。従って、接続/切断の処理は必要がない。

7.1.3 CCITT 勧告

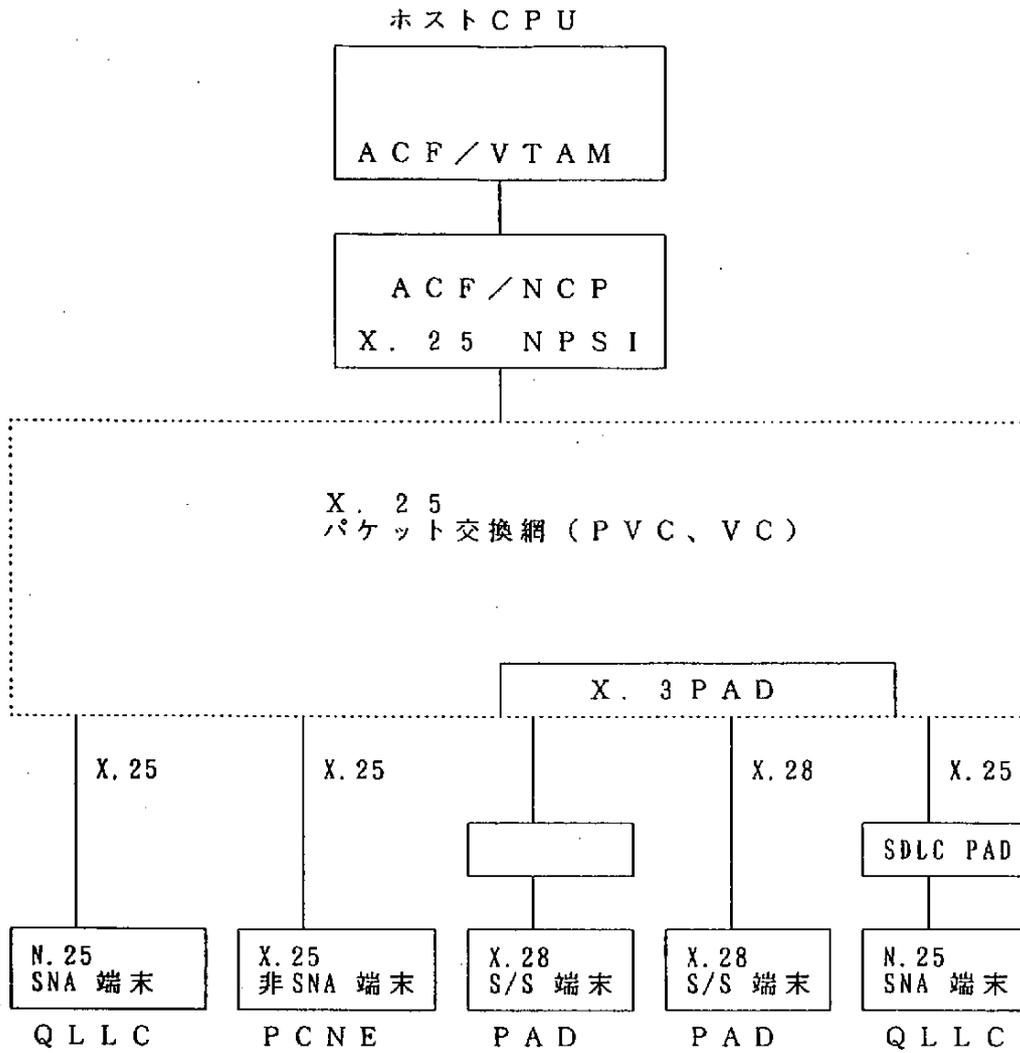
CCITT 勧告では、X.25 パケットに関連した X シリーズとして次の勧告がされている。

- X.3 : X.28 端末をパケット網と接続可能にするための PAD を規定したもの。
- X.25 : X.25 をパケット端末 (DTE) と X.25 パケット網 (DCE) とのインターフェースを規定したもの。
- X.28 : スタート・ストップ端末と X.3 のインターフェースを規定したもの。
- X.29 : X.25 パケット端末と X.3PAD のインターフェースを規定したもの。
- X.31 : ISDN 網での X.25 端末のサポートを規定したもの。
- X.32 : 公衆電話網、回線交換網を経由して、X.25 パケット網と通信を行うパケット端末とのインターフェースを規定したもの。
- X.75 : X.25 パケット交換機間の接続信号方式を規定したもの。



7.1.4 X.25NPSI による接続

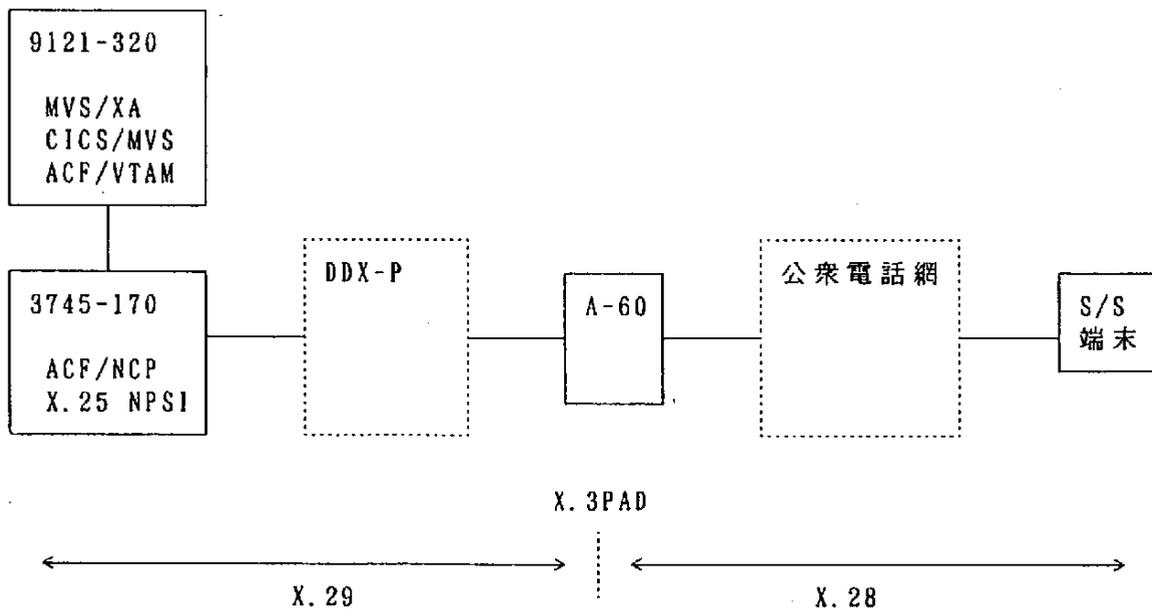
X.25 NPSI を使用したパケット網との接続では、以下のような形態をとることができる。



QLLC : Qualified Logical Link Control
 PCNE : Protocol Converter for Non-SNA Equipment
 PAD : Packet Assembly Deassembly

7.2 コミネット仙台との接続

7.2.1 接続形態



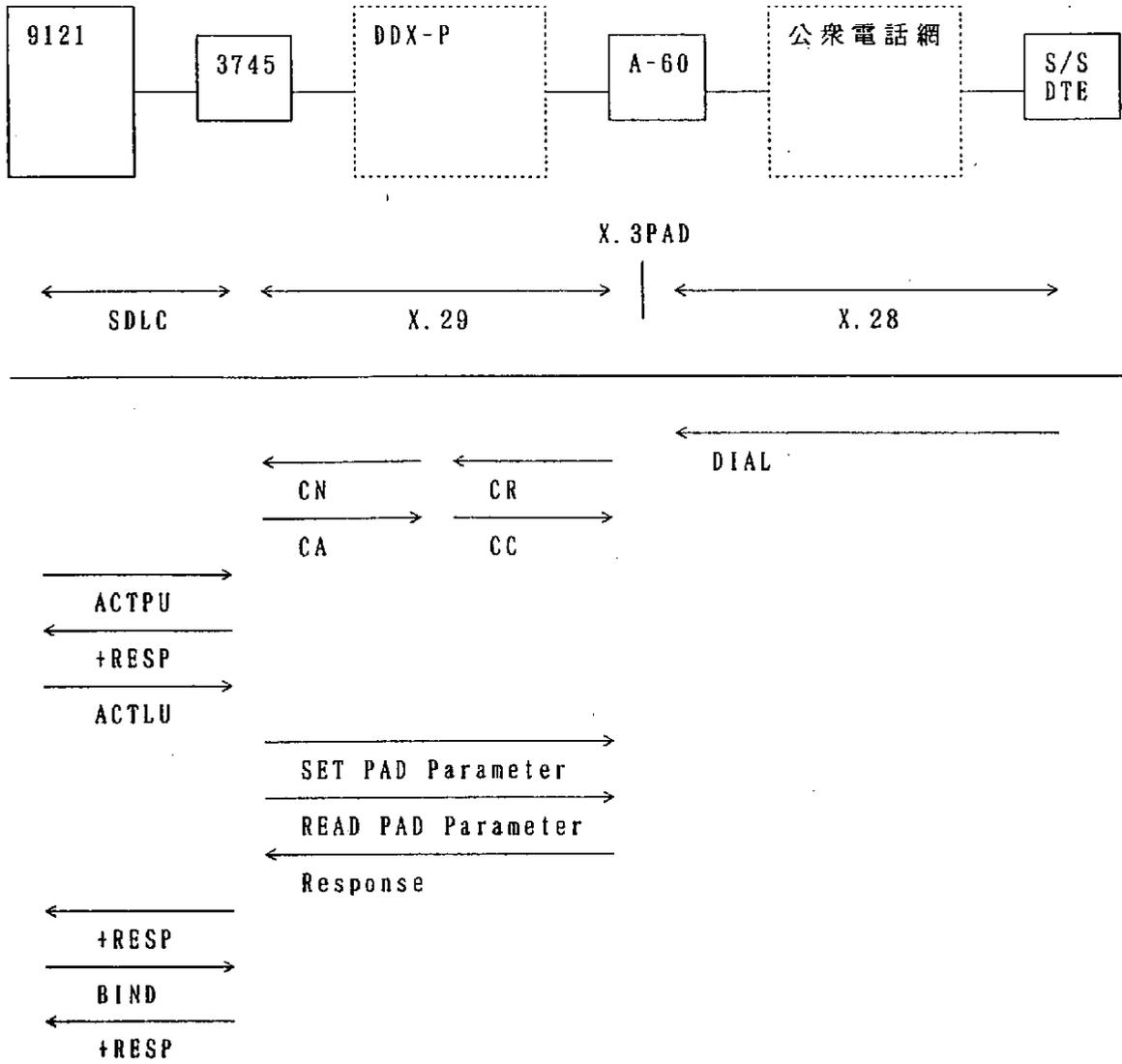
ホストのアプリケーション LU(CICS)と、NPSIの内部に作成される LU との間で LU-LU セッションがとられる。

ホスト側から見た場合、NPSI 内部に作成される LU は LU タイプ 1 のタイプライター端末になる。

NPSI と PAD (A-60) とのあいだは、X.29 のプロトコルによる接続が、PAD と S/S 端末との間は、X.28 による接続が行われる。

7.2.2 セッション確立

ホストと S/S 端末との間のセッションの確立は、次のような手順で行われる。



C R : Call Request (接続要求)

C N : Incoming Call (着呼)

C A : Call Accepted (着呼受付)

C C : Call Connected (接続完了)

7.3 資源定義

7.3.1 NCP/NPSI

X.25 パケット交換網経由で接続される資源は、X.25NPSI を使用して NCP に定義する。つまり、NPSI の提供するマクロを NCP ソースに組み込み、他の NCP 資源定義とともに生成する。

NPSI に定義された PU、LU は、NCP・NPSI がロードされる通信制御装置に存在し、SNA の資源をシミュレートする。

従って、NPSI に定義された LU とセッションを行うアプリケーションは、実際に接続される相手にかかわらず SNA 端末と見做すことができる。

7.3.2 VTAM

X.25 パケット交換網は交換回線接続となるので、NCP への定義の他に、交換大ノードの定義が必要になる。

ここに定義される資源名が実際に CICS 等の VTAM アプリケーションから参照される。

7.3.3 CICS TCT

CICS の TCT 上では、端末タイプは 3767 として定義する。

3767 は、LU タイプ 1 のタイプライター型の端末の定義である。

・交換大ノード定義例

```
SWNET  VBUILD TYPE=SWNET
PU001  PU  ADDR=01,
        DISCNT=YES,
        DLOGMOD=MODPAD,
        IDBLK=065,
        IDNUM=nnnnnn,
        MODETAB=MODETBL,
        PACING=1,
        PUTYPE=1,
        SSCPFM=USSNTO
LU00101 LU  ADDR=00
        TERM=TWX
```

・ログモード定義例

```
MODPAD  MODEENT LOGMODE=MODPAD,
```

```

FMPROF=X'03',TSPROF=X'03',
PRIPROT=X'B1',SECPROT=X'91',
COMPROT=X'3040',
PSNDPAC=1,DCODE=X'80',
PSERVIC=X'010000000000000000000000'

```

・CICS TCT 定義例

```

PAD1  DFHTCT TYPE=TERMINAL,
      ACCMETH=VTAM,
      TRMTYPE=3767,
      NETNAME=LU00101,
      TRMSTAT=TRANSCEIVE,
      TIOAL=256,
      BUFFER=256,
      RUSIZE=256,
      BRACKET=YES,
      PGESTAT=PAGE,
      PGESIZE=(12,80)

```

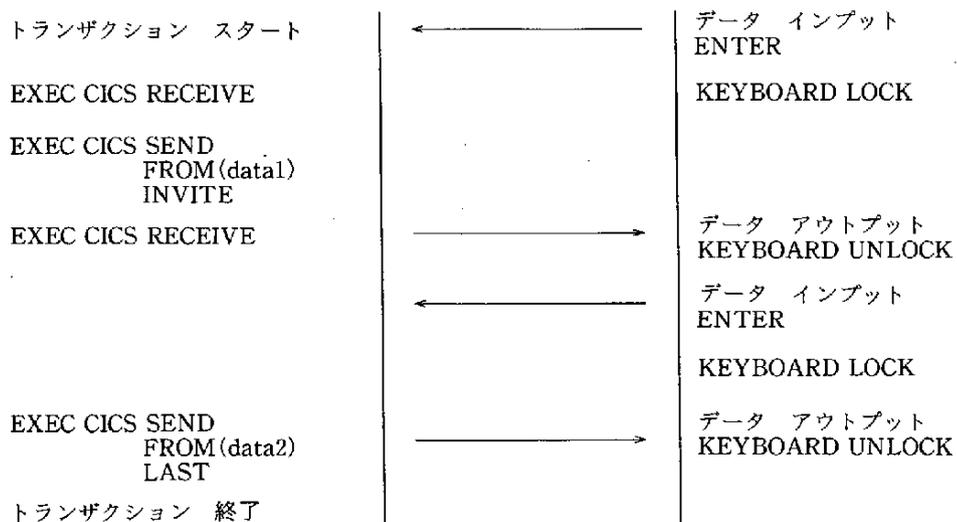
7.4 アプリケーション・プログラミング

7.4.1 概要

X.25 パケット網経由で PAD 接続された S/S 端末は、LU タイプ 1 の LU をシミュレートする。

CICS では、3767 として定義された FLIP-FLOP モードで動く。

CICS トランザクションと FLIP-FLOP モードの端末との間のデータのやりとりは、CD インディケータ (Change Direction) のセットにより行われる。



7.4.2 端末とのセッションで使用するコマンド

3767 FLIP-FLOP モードの端末とのセッションを行う時に使用するコマンドは次のようなコマンドである。

(1) RECEIVE コマンド

RECEIVE

{INTO(data-area—SET(ptr-ref))

LENGTH(data-area)

[MAXLENGTH [(data-value)]]

[NOTRUNCATE]

CONDITIONS

EOC,LENGERR,SIGNAL,TERMERR

- (a) INTO または SET を用いてデータを読み込むデータ域を指定する。
- (b) MAXLENGTH は、プログラムが受け入れるデータの最大長を指定できる。
- (c) EOC コンディションは、RU チェーンの終わりを表す。
- (d) SIGNAL コンディションは、セッションの相手から SIGNAL データフローを受け取った時に起こる。

(2) SEND コマンド

SEND

FROM(data-area)

LENGTH(data-value)

[DEST(name)]

[WAIT]

[INVITE—LAST]

[CNOTOCOMPL—DEFRESP]

CONDITIONS

SIGNAL,TERMERR

- (a) WAIT オプションを指定しないかぎりデータの送出しは、次の事象が発生するまで控えられる。
- (b) INVITE オプションを指定すると次の端末制御コマンドが RECEIVE であることを表す。つまり、CD インディケータがセットされることになる。
- (c) LAST オプションは、これがトランザクションの最後の出力操作であることを表す。

- (d) CNOTCOMPL オプションは、このコマンドで送られる RU チェーンを完了しないことを表す。
- (e) ATI トランザクションの場合、SEND コマンドはつねにトランザクション内の最初の RECEIVE コマンドの前になければならない。

(3) CONVERSE コマンド

CONVERSE
 FROM(data-area)
 FROMLENGTH(data-value)
 [INTO(data-are)—SET(ptr-ref)]
 TOLENGTH(data-area)
 [DEST(name)]
 [DEFRESP]
 [MAXLENGTH[data-value]]
 [NOTTRUNCATE]
 CONDITIONS
 EOC,LENGERR,SIGNAL,TERMERR

(4) WAIT SIGNAL コマンド

WAIT SIGNAL
 CONDITIONS
 SIGNAL

- (a) SIGNAL 条件が発生するまでタスクを一時停止させる。
- (b) HANDLE CONDITION SIGNAL コマンドを指定すればアテンションを受け取った時に該当のルーチンへブランチすることができる。

(5) ISSUE DISCONNECT コマンド

ISSUE DISCONNECT
 CONDITIONS
 SIGNAL

- (a) CICS と端末との間のセッションを終了させることができる。
- (b) ただし、TCT に
 RELREQ=(,YES)
 が指定されている場合に限る。

——禁無断転載——

平成4年3月発行

発行 財団法人 データベース振興センター
東京都港区浜松町二丁目4番1号
世界貿易センタービル7階
TEL 03-3459-8581

委託先 株式会社 河北新報社
仙台市青葉区五橋一丁目2番28号
TEL 022-211-1166

印刷所 東北堂印刷株式会社
TEL 022-245-0229

