

解説



オフィス文書の標準化と 文書データベースの研究動向†

佐藤 和 洋^{††} 絹 川 博 之^{††} 大 町 一 彦^{††}

1. はじめに

ハードウェア技術、データ処理技術、及びデータ通信技術などの進歩により高度情報処理システムが構築されつつある。低価格な高級（高機能、高性能）ワークステーション、ローカルエリアネットワークに代表される高度通信処理システムなどの出現がそれを可能にしつつあり、インテリジェントビルディング、インテリジェントシティなどの言葉も生まれている（インテリジェントネーションという言葉も出現するかもしれない）。

また、ユーザニーズの高度化及び多様化にともなって、計算機システムで扱う情報も多様化し、従来の文字／数値に加えて、テキスト、図形、画像、音声などからなる情報（これは、マルチメディアデータ、あるいは多元情報と呼ばれている。以下ではマルチメディアデータと呼ぶ。）を統一的かつ効率的に処理したいという要求が高まってきている。このような要求はオフィス⁶³⁾やエンジニアリング³⁶⁾などの情報処理分野で特に顕著である。マルチメディアデータを構成する個々のタイプのデータ（これをメディアと呼び、テキストメディア、図形メディアなどという。）に対するデータ処理技術（入出力技術も含む）の進歩と、これらのメディアを記憶格納できる各種記憶装置の出現などにより上記のような要求の実現性も高く、現在、これらのデータも蓄積されつつある。しかしながら、上記メディアデータのデータベース化（データベース管理の意味）はいまだプリミティブな形態でしか実現されておらず、国内外の大学あるいは計算機関連会社で活発な研究開発が行われている段階である⁶⁵⁾。特に、オフィス情報処理分野ではオフィス情報の国際標

準化活動^{13)~11)}、オフィス情報の電子化の高まり^{20), 21)}、OA 機器の高級化⁶²⁾などから、マルチメディアデータとしてのオフィス情報の効率的な管理及び流通システムの実用化に向けての研究開発が進められている^{23), 24), 27)~30), 34)~40), 45), 46), 48)~78)}。

本解説ではオフィス文書を対象に、オフィス文書の標準化動向とマルチメディアデータベースとしての文書データベースの研究開発動向について記述する。まず、2. ではオフィス文書の標準化動向を ISO における内容を中心に述べ、3. ではオフィス文書データベースに関する研究開発動向について述べる。具体的には、オフィス文書データモデル、ユーザインタフェース、文書データベース管理、文書データの高速処理技術などについて記述する。

2. オフィス文書の標準化動向

2.1 概 要

オフィス文書の標準化作業は下記機関で進められてきた^{13)~11)}：

- (1) ECMA
- (2) CCITT
- (3) ISO

標準化作業は ECMA が先行して実施してきており、CCITT 及び ISO での標準化作業に大きな影響を与えてきた。ECMA はオフィス文書アーキテクチャの概念確立の先駆者の役割を果たしてきたといえる。

1981年、ECMA はオフィス文書の作成、処理、交換に関する標準化を検討するための技術委員会 TC 29 を設立し、オフィス文書の標準化作業を推進してきた。そして、1985年6月には標準仕様として ECMA-101 を設定した。この間、ECMA の標準草案は CCITT や ISO における標準化作業に強く影響を与えるとともに、また逆に、これらの機関からも影響を受けている。たとえば、1984年3月に CCITT が採択した“テレマティークサービスのための文書交換プロ

† Trends of Office Document Standardization and Document Database Researches by Kazuhiro SATOH, Hiroshi KINUKAWA and Kazuhiko OHMACHI (Systems Development Laboratory, Hitachi, Ltd.).

†† 日立製作所システム開発研究所

トコル¹⁾はテキストを含む文書の交換が可能となっているが、そこで扱われるオフィス文書アーキテクチャやそのフォーマットは、ECMA が初期の段階で提案したオフィス文書に関する内容の影響を強く受けたものである。また、最近では ECMA の委員会のメンバと ISO の委員会のメンバがほとんど同じということで、ISO で進められている標準化作業は ECMA-101 標準³⁾をベースとしたものとなっている²⁾。なお、ISO では TC 97/SC 18 の WG 3 及び WG 5 で“テキスト構造”の標準化作業を推進中である。現在、DIS (Draft International Standard) 8613 として報告されており、昭和 62 年 7 月には IS (International Standards) になる見込みである。そこでは文書交換のための論理体系を規定する

オフィス文書アーキテクチャや種々のコンテンツアーキテクチャの定義及び構文体系を規定する交換フォーマットなどが定義されている⁴⁾⁻¹¹⁾。以下、本章では、ISO での標準化作業報告である DIS 8613 の概要について紹介する。他の標準化内容については文献 1), 3) を参照されたい。

2.2 ISO/DIS 8613 の概要

ISO/DIS 8613 の本題は以下になっている：“Information Processing-Text and Office Systems: Office Document Architecture (ODA) and Interchange Format (ODIF)”

ISO の標準草案は複数の Part から構成されており、

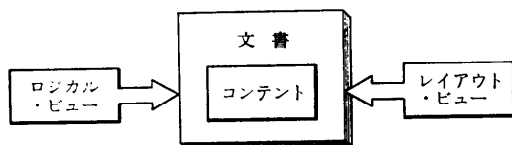


図-1 文書に対するビュー⁸⁾
(出典: ISO/DIS 8613/2 p. 11)

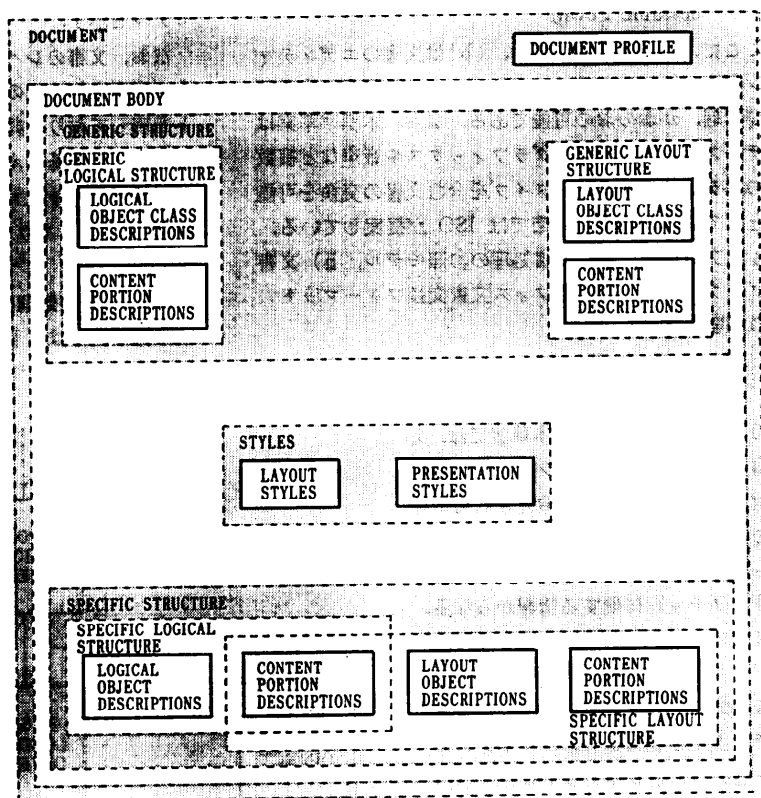


図-2 文書の記述モデル⁹⁾
(出典: ISO/DIS 8613/2 p. 20)

現在以下の 8 つの Part からなっており、その中の Part 1 から Part 6 までが DIS となっている：

- *Part 1: General Introduction⁴⁾
- *Part 2: Document Structures⁵⁾
- *Part 3: Document Processing Reference Model⁶⁾
- *Part 4: Document Profile⁷⁾
- *Part 5: Office Document Interchange Format⁸⁾
- *Part 6: Character Content Architectures⁹⁾
- *Part 7: Raster Graphics Content Architectures¹⁰⁾
- *Part 8: Geometric Graphics Content Architecture¹¹⁾

本標準草案は、i) 作成された文書の表示、ii) 文書のエディティングやフォーマット処理、を可能とするための文書の交換方法を規定している。交換における文書構造は以下のような形態をとることができる：

- (a) 書式付きフォーム (Formatted Form)
- (b) 処理可能フォーム (Processable Form)
- (c) 書式付き処理可能フォーム (Formatted Pro-

cessable Form)

ここで、(a)は文書の表示、(b)は文書のエディティングやフォーマット処理、(c)は文書の表示及び処理、がおのの可能である。なお、本標準草案はテキスト、イメージ、グラフィックスや音声など複数のメディアコンテンツタイプを含む文書の交換を可能としている。以下、本節では ISO が提案している、i) 文書構造、ii) 文書処理の参照モデル、iii) 文書プロフィール、iv) オフィス文書交換フォーマット、の概要を述べる。

2.2.1 文書構造⁹⁾

文書構造は、Part 2 (ISO/DIS 8613/2) に記述されている。本草案では、文書とはアプリケーション間で一つの単位として交換される構造化された情報の集まりで、構造(文書アーキテクチャ)及びコンテンツ(コンテンツアーキテクチャ)に関する情報からなる。文書のコンテンツビューとしては、ロジカルビュー(章、節、項などのような要素にコンテンツを関連づける見方で、上記の要素をロジカルオブジェクトという)とレイアウトビュー(ページ、エリアなど表現メディアに関連している要素とコンテンツを関連づける見方で、上記の要素をレイアウトオブジェクトという)がある(図-1)。本草案では前者をロジカル構造、後者をレイアウト構造といい、おのの文書に対する共有(generic)及び固有(specific)な構造が規定されている。

文書の構成要素としては以下のものが定義されている：

- ① 文書プロフィール
- ② ロジカルオブジェクトクラス記述
- ③ レイアウトオブジェクトクラス記述
- ④ ロジカルオブジェクト記述
- ⑤ レイアウトオブジェクト記述
- ⑥ コンテント部記述
- ⑦ 表示スタイル
- ⑧ レイアウトスタイル

図-2 上記構成要素に基づいた文書

モデルを示す。

なお、文書のレイアウト及びロジカル構造は階層構造で表現され、その各ノードは文書構造のオブジェクトを表しており、非終端ノードを複合オブジェクト、終端ノードを基本オブジェクトという。参考のために、レイアウトとロジカルオブジェクト間の関係を図-3 に、レイアウト構造の例を図-4 に、おのの示す。

2.2.2 文書処理の参照モデル⁹⁾

文書処理には、i) 編集、ii) レイアウト、iii) イメージング、の三つのプロセスがある。i) ではコン

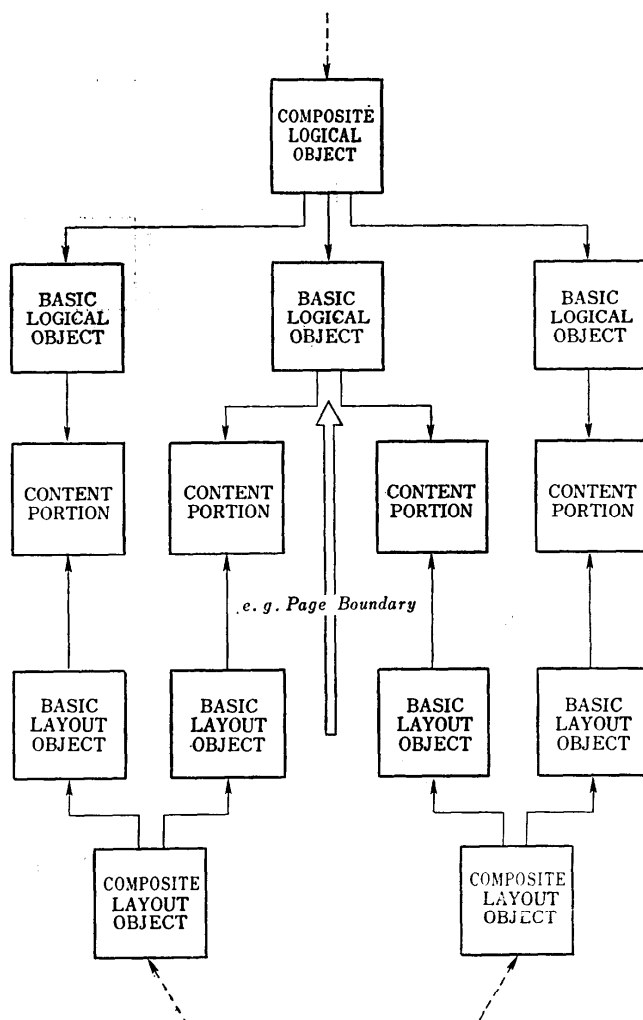


図-3 ロジカルオブジェクトとレイアウトオブジェクト間の関係⁹⁾
(出典：ISO/DIS 8613/2 p 26)

テント編集プロセスとロジカル構造編集プロセスからなり、固有ロジカル構造を生成し、ii) では文書レイアウトプロセスとコンテンツレイアウトプロセスからなり、固有レイアウト構造を生成し、iii) ではii) で得られた情報、共有レイアウト構造や表示スタイルの情報をを用いて表示媒体に文書を表示する処理からなる。これらのプロセスからなる文書処理モデルの形態を図-5 に示す。

2.2.3 文書プロフィール⁷⁾

文書プロフィールは Part 4 (ISO/DIS 8613/4) に記述されている。文書プロフィールは複数の、必須あるいはそうでない、属性で構成されており、その内容は i) 文書の構造（レイアウト／ロジカルの構造、レイアウト／表示の形式など）に関するもの、ii) 文書の特徴（文書アーキテクチャ／交換フォーマット／文書プロフィールのレベル、コンテンツアーキテクチャ、アプリケーションプロフィールなど）に関するもの、そして iii) 文書の管理（文書記述、各種日付け、作成者など）に関するもの、である。本情報は文書データのアクセスにおいて重要な役割を果たす。

2.2.4 オフィス文書交換フォーマット⁸⁾

オフィス文書交換フォーマットは Part 5 (ISO/DIS 8613/5) に記述されており、文書を交換するのに使用するデータストリームのフォーマットである。このデータストリームは“交換データユニット”と呼ばれるもので記述される。交換データユニットには以下のものがある：

- 文書プロフィール記述子；
- レイアウトオブジェクト記述子；
- レイアウトオブジェクトクラス記述子；
- ロジカルオブジェクト記述子；
- ロジカルオブジェクトクラス記述子；
- 表示形式記述子；
- レイアウト形式記述子；
- SGML エンティティ記述子；
- テキストユニット。

データストリーム内での上記交換データユニットの配置順序の規則に基づいて、本標準草案では4つのデータストリームが定義されており、おのおのデータストリーム A1, データストリーム A2, データストリーム B1, データストリーム B2 という。また、コンFORMANCEレベルに関する交換フォーマットレベルとしては、A1(データストリーム A1), A2(データストリーム A2), B(データストリーム B1 あるいは

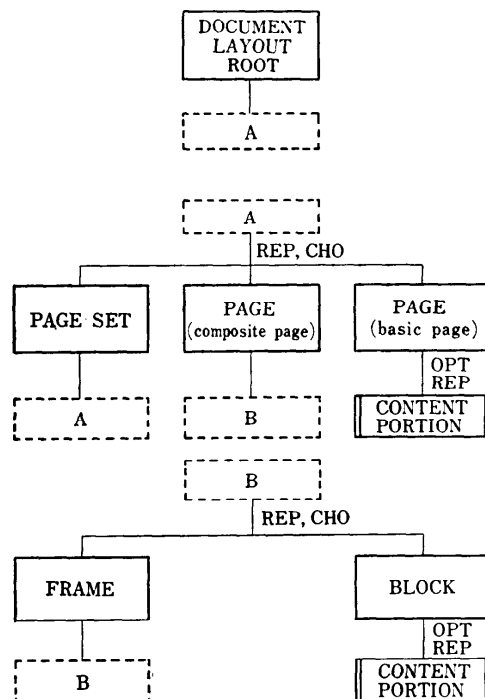


図-4 固有レイアウト構造の例⁹⁾
(出典：ISO/DIS 8613/2 p. 28)

は B2) の3レベルが定義されている。データストリームやコンFORMANCEレベルの種別は文書プロフィール記述子に指定する。

2.3 標準化活動の影響

OA 化が進むにつれ、オフィス情報の電子化が活発化してきており、計算機関連会社の各社からオフィスシステム情報の仕様体系として文書の交換アーキテクチャ (ISO の ODIF に相当) 及び内容アーキテクチャ (ISO の ODA に相当) が発表されており、文書の分散処理も可能となってきた。表現の相違はあれ、その内容は各社ともほとんど類似しており、メーリングサービス、ファイリングサービス、アプリケーションサービスなどの機能が提供されている。ちなみに、計算機関連会社の上記 ODIF/ODA に対応する例はおのおの次のようである：NTT のオフィス情報体系¹⁴⁾、IBM 社が DIA/DCA¹²⁾、日本電気(株)が IIA/ICA¹³⁾、富士通(株)が SIA/SCA¹⁵⁾、三菱電機(株)が GIA/GCA¹⁶⁾、(株)日立が MIA/MCA¹⁷⁾。上記のほかの会社も同様な情報体系を検討しているものと思われる。これらに基づいたネットワーク上での文書データベース管理システムの構築に当たっては、文書データ

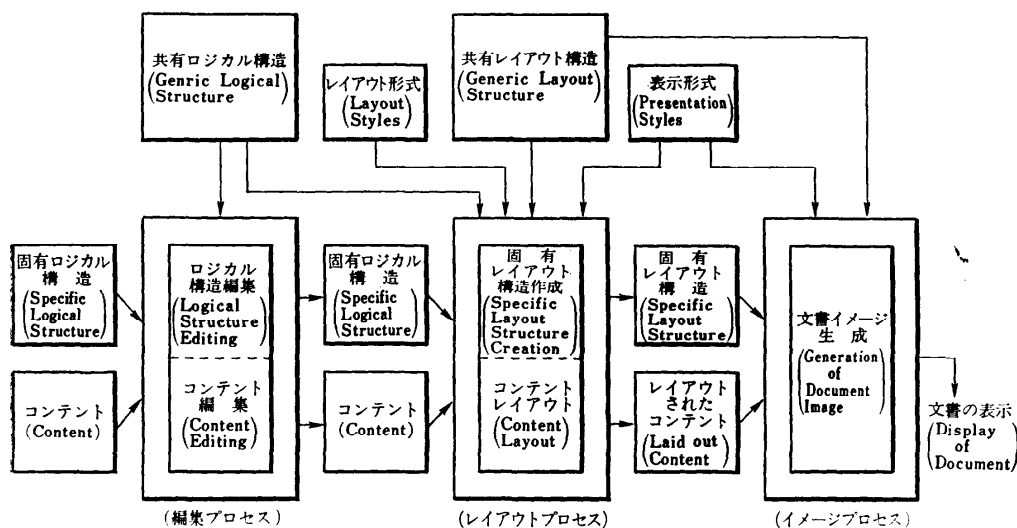


図-5 文書処理モデル⁴⁾
(出典: ISO/DIS 8613/3 p. 4)

ベース固有の課題（たとえばユーザインタフェースやアクセス法など）もあるが分散データベースに対する実用的な処理技術の提供が望まれる^{53), 56)}。

なお、上記の情報体系は OSI のアプリケーション層に位置づくものであるが、今後の情報ネットワークの急速な進展を考えるとプロトコル上の不一致はすぐに大きな問題となろう。日本語文書固有の表現形態を考慮する必要があるが⁴⁷⁾、国際的な情報ネットワークを構築、あるいはそれに参画していくには、上述した国際標準案の適用を考えていく必要がある^{2), 51)}。上記 ODA/ODIF ベースの実装規約を策定して分散データベース管理システムの構築作業を推進している INTAP の動向も注目したい。

3. オフィス文書データベースの研究動向

3.1 概 要

本章では下記事項について記述し、オフィス文書データベース研究動向を概観する：

- (1) オフィス文書データモデル
- (2) オフィス手続きとユーザインタフェース
- (3) オフィス文書データベース管理
- (4) 文書データのアクセス法とハードウェア支援

3.2 オフィス文書データモデル^{29), 36), 40), 51), 56), 63), 67), 69), 70), 74)~76)}

これまでさまざまなデータモデルが提案されてきたが、オフィスデータの特性（マルチメディアデータで

あること、複合オブジェクトであること、オブジェクト間の関連がダイナミックに変化し、その意味関連の管理を要することなど）を考えると従来のデータモデルをそのまま適用することはできない。現段階ではいまだ確定的な文書データモデルはないが、オフィス情報システム (OIS: Office Information System) における文書データのモデル化に当たっては以下のような事項を考慮すべきであろう：

(a) オブジェクト指向であること。

文書は複数の操作体系が異なるメディアデータからなっている。その中の個々メディアデータはある意味で独立しており、それらの間の関係付けが与えられて一つの文書が構成される。ここで、個々のメディアデータあるいはそれらのグループをオブジェクトとして捉えると便利である。現在、オブジェクトモデルが文書処理において有益なフレームワークを提供しているという観点から種々の提案がなされている。

(b) 抽象化機構を装備していること。

オブジェクト指向で文書を捕らえる場合、オブジェクトに関する classification, generalization/specification, aggregation などの抽象化機構¹⁸⁾が文書処理にも有用である。マルチメディアデータを操作するには必須のものと考えられる。

(c) 意味的な完全性制約が指定できること。

文書内オブジェクト間には種々の関係が存在し、これらの間の完全性を保持するためには種々の完全性制

約 (Integrity Constraints) が必要である。これには、タイプ制約, ユニーク性制約, 従属性制約, 基数制約などがあり, 指定方式としては静的な方式 (state snapshot approach) と動的な方式 (state transition approach) がある。

(d) オブジェクトの移動操作ができること。

文書は複数のオブジェクトからなり, これらのオブジェクトを再構成あるいはこれらから抽出して新たな文書を構築したりする場合, 当該機能は必須である。

(e) マルチメディアデータが扱えること。

オフィス文書にはテキスト, イメージ, 図形, 音声などさまざまな形式のデータが存在するため, これらに対する操作系が必要である。

(f) 外部表現 (論理/レイアウト構造) と内部表現

オフィス文書処理においてはユーザ指向の表現形態及び処理指向の表現形態という外部/内部表現があるが, これらの設定には標準化動向も考慮する必要がある。

以上のような観点からオフィス文書のモデル化には, 関数的データモデルや意味的 (階層) データモデルなどをベースとした提案が多くなされ, 活発な研究がなされている。

3.3 オフィス手続きとユーザインタフェース¹⁹⁾

23), 27), 28), 30), 32), 34), 36), 42), 45), 55), 63), 71)

OA は計算機及び通信装置を用いてオフィスにおける処理手続きを自動的に処理することを狙っている。したがって, このオフィス手続きを自動化するためにはオフィス手続きを十分に理解し, その仕様を明確にし, プログラムに変換する必要がある。しかしながら, 手続きの豊富さ, 例外処理の多さ, 意志決定指向的であること, さらにそれはそれが局所的であることなどから, オフィス手続きを明確な仕様で定義することは非常に困難である。実現環境とユーザ環境を考慮して対象を限定する必要がある。トロント大⁶³⁾ではフォーム処理のためのオフィス手続き仕様化機構をオフィスフォーム管理システム OFS^{23), 39)} 上に実現している。そこでは2種類の手続き, 問合せ処理とフォーム統合化処理を仕様化している。そのユーザインタフェースは TLA (Three Letter Acronym) で, フォームをオブジェクト (スケッチという) として操作する。これまで提案されている多くの OA システムは SBA¹⁹⁾, OBE²⁸⁾ や Officetalk³⁵⁾ から強く影響をうけているが TLA もその一つである。

さて, ユーザインタフェースは計算機システムにおいて最も重要なものの一つであるとともに, 最も困難なものの一つでもある。ユーザと計算機システム間には多層レベルのインタフェースが考えられ, また, 個々の実現形態も種々あり, これらをいかに実現しユーザに提供するかが大きな課題である。すなわち, ユーザインタフェースの設計においては, i) 効果的な入出力装置, ii) 簡易な構文でかつ適切な機能と量を備えた操作コマンド群, iii) 複数種類の操作機構 (直接操作あるいは記述的操作), iv) システムとの種々の対話機構 (タイプイン, ファンクションキー, メニュー, アイコン, ジェスチャ, 音声入力など), v) 容易なモードやインタフェース切替 (あるいは切替操作を意識させない), vi) フィードバック機構, vii) その他ユーザ支援ツール類などを考慮する必要がある。オフィス文書などに対するマルチメディアデータ操作においてはあいまいさのある Content Addressability を効果的にユーザインタフェース及びシステムで実現 (高速ブラウジング機構など) する必要がある, 上記事項は特に重要である。既存データベース言語の拡張³⁸⁾や種類の視覚言語^{36), 42)}の提案があるが, これからは現在注目されている知識処理を採り入れた認知科学的なアプローチが要求されよう。

3.4 オフィス文書データベース管理^{25), 36), 37), 39), 52)~54), 57), 61)~63), 65), 66), 68), 74), 75)}

文書管理はオフィスにおける最も主要な作業の一つである。現在, オフィスシステムにおける操作, たとえば, 文書のエディティング, フォーマット, ファイリング, 検索, メーリングなどの機能, 及びこれらを統合した文書データベース管理システムについて活発な研究開発が行われている。この例としてはトロント大の Officeaid^{52), 63)}, ダブリン大の OFF^{69)~73)}, ウォータールー大の MINOS⁷⁴⁾ などがある。これらを含めて現在提案されているシステムはほとんどのものがオブジェクト指向であることである。なお, 文書データベース管理システム構築のアプローチとしては^{76), 78)}, ①既存汎用 DBMS の機能拡張方式, ②新規 DBMS 方式, ③汎用 DBMS とメディアデータ専用 DBMS 群の複合方式などが考えられるが, 現状では①, ③のアプローチで研究開発が行われている。本稿ではその一つである Officeaid^{52), 63)} について紹介する。

Officeaid はオブジェクト指向に基づいたシステムであり, 利用可能なシステム資源を表す共有オブジェ

クトとして、文書、ファイルフォルダ、封筒、端末、ファイルキャビネ、プリンタ、メールトレイ、屑籠などが定義されている。オブジェクトは機能性やカテゴリによってグループ化されており、カテゴリ（クラスに対応）としては i)メタタイプ、ii)タイプ、iii)インスタンスがある。現在実現されている Officeaid はテキストのみの文書をサポートしているが、ほかのデータタイプのサポートも予定されている。文書タイプ及びインスタンスの表現には外部表現と内部表現があるが、Officeaid では文書タイプの外部表現はいくつかの文書テンプレートで定義され、このテンプレートは文書プロフィール、文書構造、文書レイアウト構造からなる。Officeaid にはこのテンプレートからの文書タイプ定義や文書タイプのビュー定義などの機能がある。また、内部表現は関係モデルのリレーションで実現している (MISTRESS リレーショナル DBMS)。

図-6 に Officeaid のシステムアーキテクチャを示す。Officeaid は UNIX* の下、C 言語で SUN ワークステーション上で実現されている。図において、デフォルト環境 (Default Environment) はすべてのユーザに共通の機能を提供し、アプリケーション固有環境 (Application-Specific Environment) は特殊ユーザ用のものである。前者においては、文書のエディティング、フォーマット、ファイリング、検索、メールリングなどの機能が提供されており、特に検索に関してはブラウジングやファジィ問合せ機構がある。後者においてはドキュメントタイプに基本と仮想の二つのタイプを設けてデータベースシステムにおけるビュー機構を提供している。

3.5 文書データのアクセス法とハードウェア支援

本節では文書データのアクセス法とハードウェア支援について概観する。メディアとして画像、図形、音声などを含んだオフィス文書に対してははまだ十分な研究が行われていないので、以下の説明では文書は属性とテキストから構成されているものとする。

オフィス環境ではデータの挿入が頻繁であり、削除や更新は少なく、検索もそれほど多くなく、ほとんどの文書はアクセスされないという特徴をもち、長大可変長データの扱いなど、従来の DBMS 適用環境とは趣きを異にしている。したがって、これまで提案され、商用 DBMS で使用されている既存のアクセス法をそのまま適用することには問題がある。ここではテキストに限定して、これに対して提案されているアク

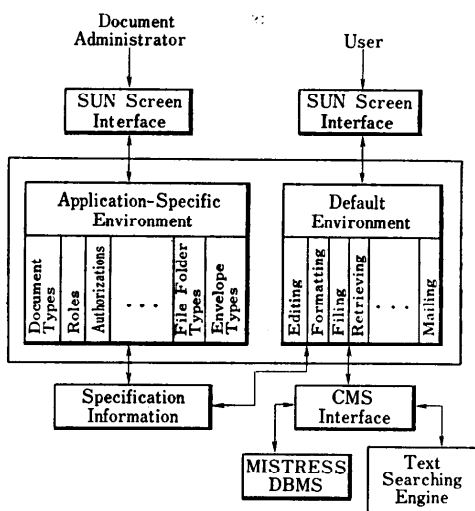


図-6 Officeaid のシステムアーキテクチャ⁴³⁾
(出典: "Office Automation", Springer-Verlag (1985) p. 29)

セス法及びそのハードウェア支援方式について記述する。

(1) 文書データ (テキスト) のアクセス法^{43), 44), 59), 63), 64), 72)}

テキスト検索に対するアクセス法としては、大きく分けて次の4つがある:

- ① 全テキスト走査法
- ② 転置ファイル法
- ③ Superimposed coding 法
- ④ クラスタリング法

①は余分なスペース負荷はなく、挿入や更新処理の負荷も小さいが、検索時間が長いという欠点があり、大規模データベースへの適用には問題がある。このためハードウェア支援方式が提案されている。また、②に関しては、転置インデックスとして各種の構成法が適用でき、検索時間は短い、インデックス用の領域を必要とし、挿入や更新の負荷が大きく、また、"Don't Care" 文字を含む探索操作が困難などの欠点がある。さらに、③は提案されたのは古いが最近また注目されてきている方法である。スペースや挿入処理負荷も少なく、データベースの成長に対して柔軟性をもつが、ベースとなる符号ファイル (Signature file) に対するアクセスがシーケンシャル探索であることから検索時間が長いという欠点がある。これに対しても専用プロセッサの提案がある。特に、トロント大学などで活発な研究が行われている^{59), 63)}。最後に④であるが、挿入

* UNIX はベル研究所の登録商標である。

時にクラスタリング処理が必要となり、動的な環境では問題があるがスペース負荷や検索時間はそれほど問題とはならない。

以上のように現在これといったアクセス法があるわけではなく、従来のものとの組合せあるいはその拡張で文書データに対するアクセス法が検討されてい

る。文書プロフィール情報など、定型的なデータ構造も文書データの一部であり、既存のアクセス法の拡張^{24), 26)}で十分対応できる点もあるが、マルチメディアデータを含む文書に対するアクセスを効率化するためには内容検索性 (Content Addressability) の実現

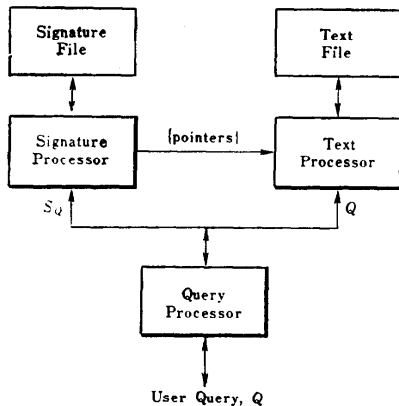


図-7 テキスト検索システム (TRS) の構成例⁴³⁾
(出典: "Office Automation", Springer-Verlag
(1985) p. 346)

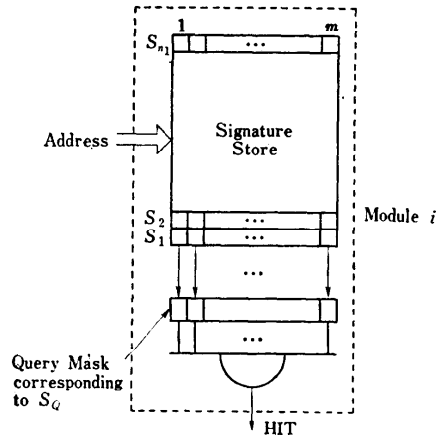


図-8 WSBP 構成の符号ファイルプロセッサの例⁴³⁾
(出典: "Office Automation", Springer-Verlag (1985)
p. 349)

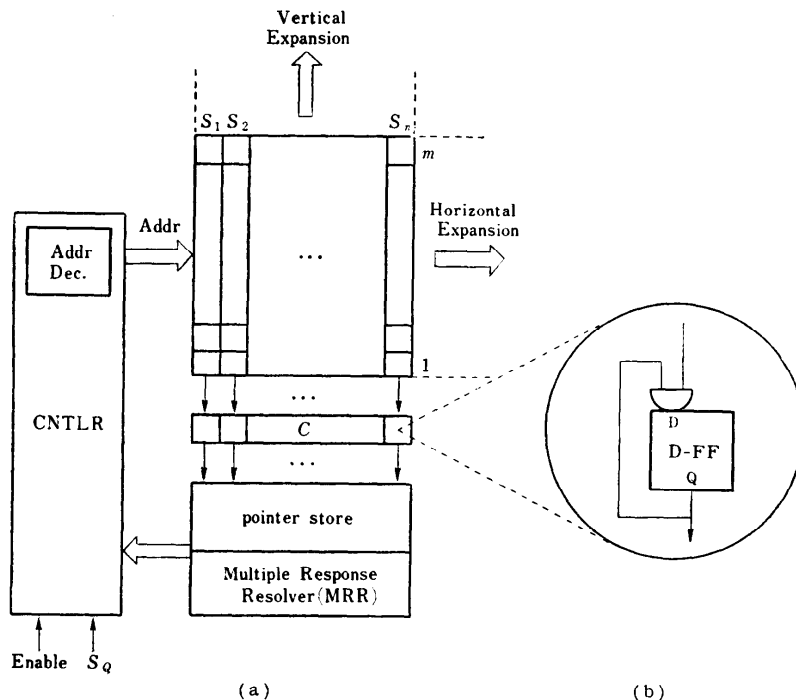


図-9 WPBS 構成の符号ファイルプロセッサ (a) とそのシリアル比較器 (b)⁴³⁾
(出典: "Office Automation", Springer-Verlag (1985) p. 351)

方式、大量な可変長データの効率的な管理方式などが大きな課題といえる^{63)~65),74)}。

(2) ハードウェア支援方式^{33),41),44),63),64)}

データベースの分野ではデータベースマシン²²⁾として活発な研究が行われているが、テキスト処理の分野でもテキスト検索マシンとして研究が行われてきている⁶³⁾。データベースマシンは、特に関係データベースにおける関係代数演算の高速処理を目的としたものであるが、テキスト検索マシンはパターンマッチングの高速処理を目的としたものである。符号ファイルをベースにしたテキスト検索システムの構成例を図-7に示す。図の符号プロセッサの構成としてはベル研究所が提案している WSBP (Word-serial, bit-parallel) 構成 (図-8 参照) とトロント大が提案している WPBS (Word-parallel, bit-serial) 構成 (図-9 参照) がある⁶³⁾。また、テキスト検索マシンはパターンマッチャであり、これは “Logic-with-pattern” と “Logic-with-text” の二つに分類される。前者の場合、処理能力は格納されているパターンの量に関係し、パターンは直接あるいはコード化されて複数のプロセッサあるいは論理セルに格納されている。一方、後者の処理能力は格納されているテキストストリングの量に関係する。前者より高速ではあるがテキストを収容するための大規模なハードウェアを必要とする。前者の例としては連想メモリ、セルラロジックアレイや有限オートマトンに基づいたものがあり、後者の例としては連想線形アレイプロセッサ、連想線形テキストプロセッサなどがある⁶³⁾。なお、上記のような多重処理システムにおいては多重応答解決方式 (MRR) が重要な課題となる。これに関しては、アドレス順検索方式や値順検索方式など、各種の提案がなされている⁶⁴⁾。

4. おわりに

オフィス文書の標準化動向とオフィス文書データベースの研究動向について概説した。電子化情報の自由な流通及び管理を可能にするためには、その流通機構の確立とともに電子化情報構造の標準化が必要である。今後の社会においては、さらに情報の流通が活発になるはずであり、その兆候はすでにパソコン通信などに現れている。情報の共有化を図るためにも情報構造や通信プロトコルの標準仕様が早期に設定される必要がある。オフィス文書は遣り取りが頻繁であり、その標準仕様の設定は有用である。ISO など各種標準化機関の動向に注目したい。また、マルチメディアデー

タベースとしての文書データベースに関しては、そのデータ特性 (多種のデータ、大量なデータ、複雑なデータ構造、など) 及び処理特性 (多様な処理手続き、統合的な処理試行錯誤の処理、複合オブジェクト処理など) からさまざまな実現アプローチが考えられ、現在活発な研究が実施されている。基本的にはセマンテックスの扱いが課題であり、これまでのデータベース研究及び知識処理研究などを総合した研究が必要である。なお、本解説では触れなかったが、文書データ (マルチメディアデータ) 通信処理技術の確立が必須である。これに関しても研究開発が活発化しており、これによって近い将来、分散データベースとしての分散文書 (マルチメディア) データベース管理システムが出現するであろう^{30),31),48)~50),53),63),66),68)}。

参 考 文 献

- 1) Document Interchange Protocol for the Tele-matic Services, CCITT Recommendation T. 73 (Geneva, Mar. 1984).
- 2) Horak, W.: Office Document Architecture and Office Document Interchange Formats: Current Status of International Standardization, Computer, pp. 50-60 (Oct. 1985).
- 3) Standard ECMA-101: Office Document Architecture, ECMA (Sep. 1985).
- 4) ISO/DIS 8613/1 Part 1: General Introduction, Information Processing-Text and Office Systems-Office Document Architecture (ODA) and Interchange Format (June 13, 1986).
- 5) ISO/DIS 8613/2, Part 2: Document Structures, Information Processing-Text and Office Systems-Office Document Architecture (ODA) and Interchange Format (Sep. 13, 1986).
- 6) ISO/DIS 8613/3, Part 3: Document Processing Reference Model, Information Processing-Text and Office Systems-Office Document Architecture (ODA) and Interchange Format (May 1986).
- 7) ISO/DIS 8613/4, Part 4: Document Profile, Information Processing-Text and Office Systems-Office Document Architecture (ODA) and Interchange Format (1986).
- 8) ISO/DIS 8613/5, Part 5: Office Document Interchange Format (ODIF), Information Processing-Text and Office Systems-Office Document Architecture (ODA) and Interchange Format (1986).
- 9) ISO/DIS 8613/6, Part 6: Character Content Architectures, Information Processing-Text and Office Systems-Office Document Architecture (ODA) and Interchange Format (Apr.

- 1986).
- 10) ISO/DP 8613/7, Part 7: Raster Graphics Content Architectures, Information Processing-Text and Office Systems-Office Document Architecture (ODA) and Interchange Format (June 1986).
 - 11) ISO/DP 8613/8, Part 8: Geometric Graphics Content Architectures, Information Processing-Text and Office Systems-Office Document Architecture (ODA) and Interchange Format (June 1986).
 - 12) オフィスシステム体系, 日経データプロ. OA, OA 1-110-301~OA 1-110-306 (1983年9月).
 - 13) オフィス情報体系 OIA, 日経データプロ. OA, OA 1-110-761~OA 1-110-766 (1983年10月).
 - 14) 統合オフィスシステム NIS, 日経データプロ. OA, OA 1-110-771~OA 1-110-777 (1984年2月).
 - 15) FOSを支えるアーキテクチャと新しい技術, FACOM ジャーナル, Vol. 10, No. 10, pp. 99-102 (1984).
 - 16) 石田他: MELCOM 統合 OA アーキテクチャ, 三菱電機技報, Vol. 60, No. 7, pp. 447-451 (1986年).
 - 17) 笹部他: 日立におけるマイクロメインフレーム結合の設計思想, 情報処理学会「マルチメディア通信と分散処理研究会」報告 (1987年3月5日).
 - 18) Smith, J.M. et al.: Database Abstractions: Aggregation and Generalization, ACM Trans. Data Base Syst., Vol. 2, No. 2, pp. 105-133 (June 1977).
 - 19) de Jong, P.: The System for Business Automation: A Unified Application Development System, Proc. of IFIP-80, pp. 469-474 (1980).
 - 20) Morgan, H.L.: Research and Practice in Office Automation, Proc. of IFIP-80, pp. 783-789 (1980).
 - 21) Ellis, C. et al.: Computer Science and Office Information Systems, ACM Comput. Surv., Vol. 12, No. 1, pp. 27-60 (Mar. 1980).
 - 22) Hsiao, D.K.: Data Base Computers, in Advances in Computers, Vol. 19, ed. C.Y. Marshall, Academic Press, pp. 1-64 (1980).
 - 23) Tsichritzis, D.C.: OFS: An Integrated Form Management System, Proc. of the 6th VLDB Conf., pp. 161-166 (1980).
 - 24) Litwin, W.: Linear Hashing: A New Tool for File and Table Addressing, Proc. of 6th VLDB Conf., pp. 212-223 (Oct. 1980).
 - 25) McLeod, I.A.: A Data Base Management System for Document Retrieval Applications, Inf. Syst., Vol. 6, No. 2, pp. 131-137 (1981).
 - 26) Robinson, J.T.: The k-D-B-Tree: A Search Structure for Large Multidimensional Dynamic Indexes, Proc. of the ACM SIGMOD Conf., pp. 10-18 (1981).
 - 27) Luo, D. et al.: Form Operation by Example: A Language for Office Information Processing, Proc. of the ACM SIGMOD Conf., pp. 212-223 (1981).
 - 28) Zoelf, M.M.: QBE/OBE: A Language for Office and Business Automation, IEEE Computer 14, pp. 13-22 (May 1981).
 - 29) Kowarski, L. et al.: The Document Concept in a Data Base, Proc. of the ACM SIGMOD Conf., pp. 276-283 (1982).
 - 30) Rabitti, F. et al.: A Distributed Form Management System with Global Query Facilities, in Office Information Systems, ed. N. Naffah, North-Holland (1982).
 - 31) Schiker, P.: Naming and Addressing in a Computer-Based Mail Environment, IEEE Trans. Comm., Vol. COM-30, No. 1, pp. 46-62 (Jan. 1982).
 - 32) Shu, N.C. et al.: Specification of Forms Processing and Business Procedures for Office Automation, IEEE Trans. Softw. Eng., Vol. SE-8, No. 5, pp. 499-512 (Sep. 1982).
 - 33) Burkowski, F.J.: A Hardware Hashing Scheme in the Design of a Multiterm String Comparator, IEEE Trans. Comput., Vol. C-3, No. 9, pp. 825-834 (Sep. 1982).
 - 34) Furuta, R. et al.: Document Formatting Systems: Survey, Concepts, and Issues, Comput. Surv., Vol. 14, No. 3, pp. 417-473 (Sep. 1982).
 - 35) Ellis, C. et al.: Office Talk-D: An Experimental Office Information System, Proc. 1st ACM SIGOA Conf., pp. 131-140 (June 1982).
 - 36) Gardarin, G. and Gelenbe, E. (Eds): New Applications of Data Bases, Academic Press (1983).
 - 37) Economopoulos, P. et al.: A System for Managing Image Data, Proc. of IFIP-83, pp. 89-94 (1983).
 - 38) Stonebraker, M. et al.: Document Processing in a Relational Database System, ACM Trans. Office Inf. Syst., Vol. 1, No. 2, pp. 143-158 (Apr. 1983).
 - 39) Tsichritzis, D.C. et al.: A Multimedia Office Filing System, Proc. of 9th VLDB Conf., pp. 2-7 (1983).
 - 40) Gibbs, S.J. et al.: A Data Modelling Approach for Office Information Systems, ACM Trans. Office Inf. Syst., Vol. 1, No. 3, pp. 299-319 (1983).
 - 41) Haskin, R.L. et al.: Operational Characteristics of a Hardware-Based Pattern Matcher, ACM Trans. Database Syst., Vol. 8, No. 1, pp. 15-40 (Mar. 1983).

- 42) Hammer, M. et al.: What Makes a Good User Interface?, Proc. AFIPS Office Automation Conf. (Feb. 1983).
- 43) Hollaar, L. A. et al.: Architecture and Operation of a Large, Full-Text Information-Retrieval System, in Advanced Database Machine Architecture, ed. D. K. Hsiao, Prentice-Hall, pp. 256-299 (1983).
- 44) Larson, P. A.: A Method for Speeding Up Text Retrieval, Proc. of ACM SIGMOD Conf. (May 1983).
- 45) Lee, D. L. et al.: Voice Response System, ACM Comput. Surv., Vol. 15, No. 4 (Dec. 1983).
- 46) Bracchi, G. et al.: SOS: A Conceptual Model for Office Information Systems, Proc. of the ACM SIGMOD Conf., pp. 108-116 (1983).
- 47) Kinukawa, H. et al.: A Study of Japanese Document Interchange, Proc. of ICTP '83, pp. 265-269 (Oct. 1983).
- 48) Tschritzis, D. C.: Message Addressing Schemes, ACM Trans. Office Inf. Syst., Vol. 2, No. 1, pp. 58-87 (Jan. 1984).
- 49) Christodoulakis, S. et al.: Design Considerations for a Message File Server, IEEE Trans. Softw. Eng., Vol. SE-10, No. 2, pp. 201-210 (Mar. 1984).
- 50) Hogg, J. et al.: An Active Mail System, Proc. ACM SIGMOD '84 Conf., SIGMOD Record, Vol. 14, No. 2 (June 1984).
- 51) Horak, W. et al.: An Object-Oriented Office Document Architecture Model for Processing and Interchange of Documents, Proc. of the 2nd ACM SIGOA Conf. (June 1984).
- 52) Lee, A. et al.: Officeaid: An Integrated Document Management System, Proc. of ACM SIGOA Conf., pp. 170-180 (June 1984).
- 53) Lyngback, P. et al.: Object Management in Distributed Information Systems, ACM Trans. Office Inf. Syst., Vol. 2, pp. 96-122 (No. 2, 1984).
- 54) Ahlsen, M. et al.: An Architecture for Object Management in OIS, ACM Trans. Office Inf. Syst., Vol. 2, No. 3, pp. 173-196 (July 1984).
- 55) Yao, S. B. et al.: FORMANAGER: An Office Forms Management System, ACM Trans. Office Inf. Syst., Vol. 2, No. 3, pp. 235-262 (July 1984).
- 56) Zdonik, S.: Object Management System Concepts, Proc. of the 2nd ACM SIGOA Conf., pp. 13-19 (1984).
- 57) Christodoulakis, S. et al.: Development of a Multimedia Information System for an Office Environment, Proc. of 10th Conf. on VLDB, pp. 261-271 (Aug. 1984).
- 58) Andrews, H. L.: Speech Processing, IEEE Computer, pp. 315-324 (Oct. 1984).
- 59) Faloutsos, C. et al.: Signature Files: An Access Method for Documents and its Analytical Performance Evaluation, ACM Trans. Office Inf. Syst., Vol. 2, No. 4 (Oct. 1984).
- 60) Mazer, M. S. et al.: Logical Routing Specification in Office Information Systems, ACM Trans. Office Inf. Syst., Vol. 2, No. 4 (1984).
- 61) Woo, C. C. et al.: Authorizations in a Computer-based Office Information System, Proc. of the 1st IEEE OA Conf. (Dec. 1984).
- 62) Choy, D. M. et al.: A Database Management System for Office Systems and Advanced Workstations, SIGOA Newsl., Vol. 5, No. 3, pp. 17-24 (Fall 1984).
- 63) Tschritzis, D. C. (Eds): Office Automation, Springer-Verlag (1985).
- 64) Faloutsos, C.: Access Methods for Text, Comput. Surv., Vol. 17, No. 1, pp. 49-74 (Mar. 1985).
- 65) Christodoulakis, S.: Multimedia Data Base Management: Applications and Problems, Proc. of ACM SIGMOD Conf. (1985).
- 66) Poggio, A.: CCWS: A Computer-Based, Multimedia Information System, Computer, pp. 92-103 (Oct. 1985).
- 67) Peels, A. J. H. M. et al.: Document Architecture and Text Formatting, ACM Trans. Office Inf. Syst., Vol. 3, No. 4, pp. 347-369 (Oct. 1985).
- 68) Thomas, R. H. et al.: Diamond: A Multimedia Message System Built on a Distributed Architecture, Computer, pp. 65-77 (Dec. 1985).
- 69) Harper, D. J. et al.: MINSTREL-ODM: A Basic Office Data Model, Inf. Process. Manage., Vol. 22, No. 2, pp. 83-107 (1986).
- 70) Lamersdorf, W. et al.: A Recursive Approach to Office Object Modelling, Inf. Process. Manage., Vol. 22, No. 2, pp. 109-120 (1986).
- 71) Morrissey, J. M. et al.: Interactive Querying Techniques for an Office Filing Facility, Inf. Process. Manage., Vol. 22, No. 2, pp. 121-134 (1986).
- 72) Smeaton, A. F. et al.: Information Retrieval in an Office Filing and Future Work in Project MINSTREL, Inf. Process. Manage., Vol. 22, No. 2, pp. 135-149 (1986).
- 73) Restorick, F. M.: Novel Filing Systems Applicable to an Automated Office: A State-of-the-Art Study, Inf. Process. Manage., Vol. 22, No. 2, pp. 151-172 (1986).
- 74) Christodoulakis, S. et al.: The Multimedia Object Presentation Manager of MINOS: A Symmetric Approach, Proc. of ACM SIGMOD Conf., pp. 295-310 (1986).
- 75) Woelk, D. et al.: An Object-Oriented Approach to Multimedia Databases, Proc. of ACM SIGMOD Conf., pp. 311-325 (1986).
- 76) 情報処理学会「データベースシステム研究会」報告 (1985年11月18日).
- 77) データベースシステムに関する調査: (社)日本電子工業振興協会 (1985年3月).
- 78) データベースシステムに関する調査: (社)日本電子工業振興協会 (1986年3月).

(昭和62年4月16日受付)

訂 正

第 28 巻第 6 号 pp. 710～720 に掲載しました佐藤氏・他著の「オフィス文書の標準化と文書データベースの研究動向」を下表のように訂正いたします。

ページ	行	誤	正
711	左 21	昭和 62	1988
711	左 22	年 7 月には IS (Internation-	年 1 月には IS (Internation-
711	左 36	(ODIF)	削除する。
713	右 8	(ISO の ODIF に相当)	削除する。
713	右 8 } 右 9 }	内容アーキテクチャ (ISO の ODA に相当)	内容のアーキテクチャ (ISO の ODA/ODIF に相当)
713	右 14	上記 ODIF/ODA に対応する例	文書の交換及び内容のアーキテクチャ