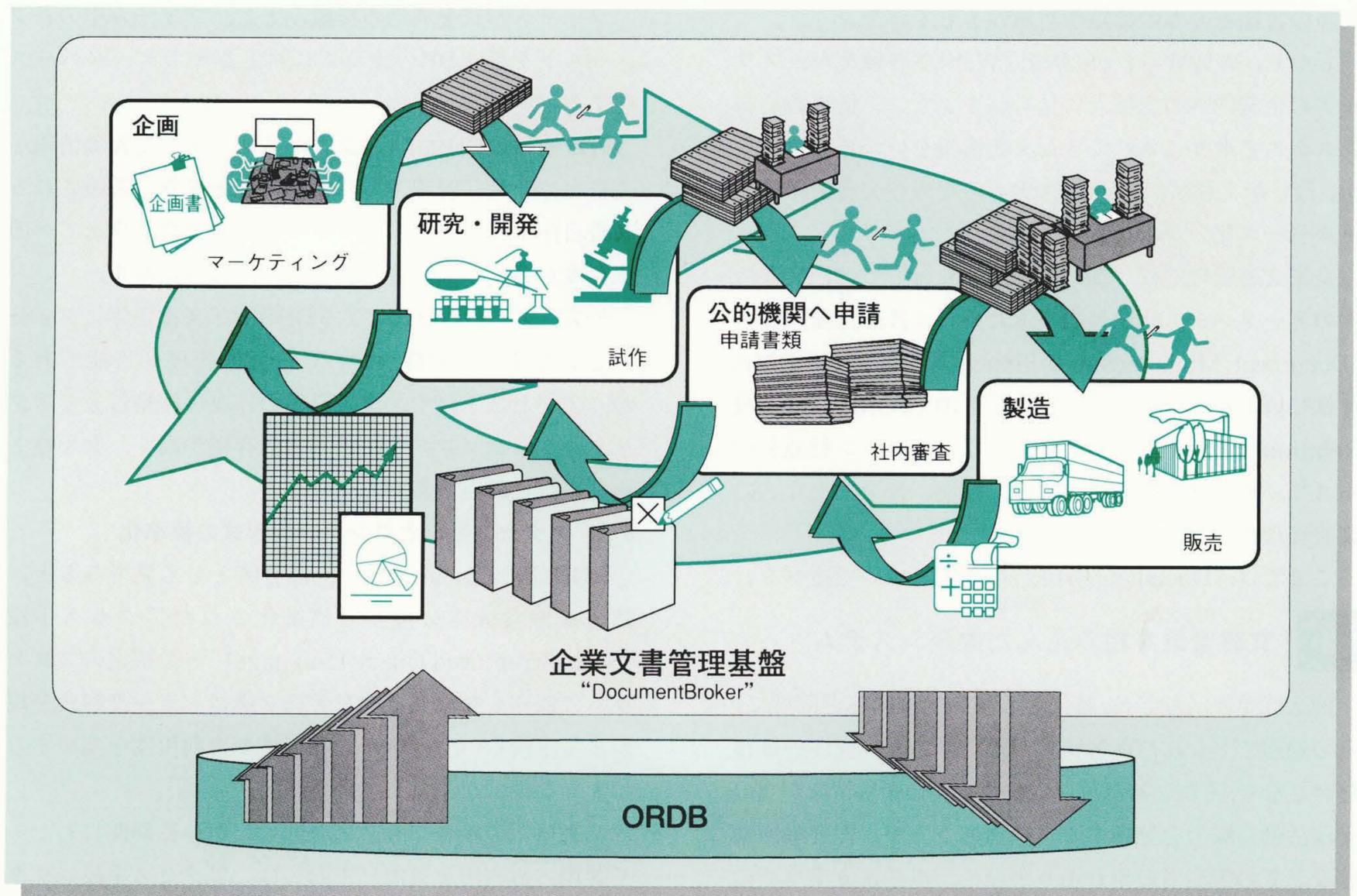


# 分散オブジェクト ベース アプリケーション 文書管理システム

## Distributed Object-Based Applications for Document Management

和歌山 哲 Satoshi Wakayama 福田俊彦 Toshihiko Fukuda  
伊藤泰樹 Yasuki Itô 加藤寛次 Kanji Katô



注：略語説明 ORDB (Object Relational Database)

### 企業文書管理基盤“DocumentBroker”を中心とした文書業務システム

企業の業務処理で、文書情報の活用を実現するためのシステムの基盤として“DocumentBroker”がある。企画から製造・販売に至る、企業の業務活動の中の「文書情報」を管理し、組織活動の中での活用を支援する業務システムが求められている。

次世代の業務システム構築に向けた基盤として、アプリケーションフレームワークの主要な柱の一つである企業文書管理基盤“DocumentBroker”を中心とした文書管理システムを開発した。

このシステムのねらいは、基幹的な業務システムの中で文書情報を扱えるようにし、それによって従来の数値系データだけでなく、文書情報を有効活用する新しい基幹業務システムを構築するための基盤を提案することにある。

企業文書管理基盤“DocumentBroker”では、文書モ

デルとAPI(Application Programming Interface)の標準化の動向を踏まえ、文書管理の業界標準化団体であるAIIM(Association for Information and Image Management International)の提唱する標準仕様DMA(Document Management Alliance)をいち早く採用し、これに構造化文書処理対応の拡張を加えた。また、文書の管理情報と文書自体の一元管理を実現するために、データベースとしてORDB(Object Relational Database)を採用し、構造化文書にも対応できる文書処理系プラグイン機能を開発した。

## 1 はじめに

従来の業務システムでは、例えば申請書類や注文書などの書類中の顧客名や商品名など、オンライン業務処理に必要な情報だけを顧客コード、商品コードなどのマスターデータとしてそれぞれ処理し、それ以外の文書情報や実際の書類そのものは処理対象外としてきた。

しかし、WWW(World Wide Web)や各種グループウェアの企業内への急拡大に見られるように、従来の業務システムで取り残されてきた文書情報をいかに活用するか、言い換えれば、「業務システムのための文書情報のデータベース化」が求められている。

企業文書管理基盤“DocumentBroker”では、文書情報のデータベース化を実現するために、標準仕様“DMA(Document Management Alliance)”を採用し、DMA仕様に構造化文書処理対応の拡張を加え、ORDB(Object Relational Database)の文書処理系プラグイン機能をベースに、構造化文書の構成管理、検索、部分抽出などの文書管理機能を提供する。

ここでは、DocumentBrokerの概要について述べる。

## 2 文書管理を組み込んだ業務システム

例えば受注活動では、結果としての受注内容と同様に、その過程でつかんだ顧客情報が重要であり、受注の推移だけでなく、それと結び付けられた折々の顧客情報が次の受注活動に威力を発揮する。同様のことが、保守整備業務などでの担当者の申し送り事項や、申請審査業務での担当者の判断根拠に関する付帯情報などにも当てはまる。

各種製造業での企画から製造・販売に至る企業活動を見てみると、おのおのの業務活動の結果・成果として最も重要な流れの一つが「文書情報」である。これまでの業務システムで扱われてきた情報と文書情報を有機的に結び付け、組織内で迅速に処理し、かつそこから有効な情報を抽出し、これを活用することのできる組織活動とそれを支援する業務システムが求められている(43ページの図参照)。

## 3 文書情報のデータベース化

### 3.1 文書情報のデータベース化の課題

ワードプロセッサなどで作成する一般的な文書も含め、広範囲かつ長期間にわたって共用、再利用するためのDB(Database)化を図るには、何らかの共通的な枠組みが必要である。従来の文書情報のDB化には、以下の二

つの大きな流れがある。

#### (1) 管理情報のDB化

文書を管理する管理情報に共通の枠組みを与え、管理情報をDB化する。文書自体はそのまま蓄積するために二元的管理となる。

#### (2) フォーム形式によるDB化

フォーム形式を共通の枠組みとし、フォーム中の各フィールドを従来DBのカラムに対応させてデータベース化する。

管理情報のDB化では、文書から切り出した管理情報と文書自体を別に管理する二元的管理となり、蓄積される文書自体を作成したアプリケーションソフトウェアへの依存度が高く、共用性や再利用性に難点がある。

一方、フォーム形式によるDB化は、文書自体にフォームという枠組みを持つために一元的管理が可能であるが、アクセス手段や、文書自体の作成・編集などでアプリケーションソフトウェアへの依存度が高く、共用性と再利用性に難点がある。

### 3.2 アクセス手段とコンテンツ形式の標準化

RDB(Relational Database)を例として見てみると、強力な検索機能を備えて標準化されたアクセス手段“SQL(Structured Query Language)”と、特定のプログラムに依存しない共通的なデータ構造・データ形式を規定する手段(スキーマ)が、共用性や再利用性を高めるのに大きな役割を果たしている。

文書情報の場合でも、広範囲に、かつ長期間にわたって共用・再利用を可能にするには、アクセス手段の標準化と、蓄積するコンテンツ形式の標準化が必要となる。

DocumentBrokerでは、AIIM(Association for Information and Image Management International)の提唱する標準仕様をいち早く採用することによってアクセス手段の標準化を目指す。また、構造化文書によるコンテンツ形式の標準化の流れに対応し、標準仕様に構造化文書処理対応の拡張を加え、構造化文書の構成管理、検索、部分抽出などの管理機能を提供していく。

### 3.3 DMA<sup>1)</sup>の仕様とDocumentBrokerでの拡張

#### 3.3.1 DMAのオブジェクトモデルとその拡張

DMAは、文書処理アプリケーションに対する統一的な文書管理API(Application Programming Interface)と文書モデルによって文書管理システムの業界標準仕様を策定することを目的とし、文書管理の業界標準化団体AIIMのタスクフォースとして1995年4月に発足し、1997年末に標準仕様“DMA 1.0”が決定された。

DMAのオブジェクトモデルは、文書を管理するオブジェクト群とコンテナ、オブジェクト間のリンクで構成され、一つの論理的文書に対して、複数のバージョン系列、複数のコンテンツ形式、複数の構成要素(ファイル)の管理が可能なモデルになっている(図1参照)。

各オブジェクトクラスは、上記の機能を実現するための基本的な役割(表1参照)を持ち、それに対応した操作機能が標準APIとして規定されている。

DocumentBrokerでは、DMAの基本仕様に対して、SGML(Standard Generalized Markup Language)構造化文書管理機能や全文検索機能を追加するためのモデルの拡張と、これに対応するAPI仕様の拡張を行っている。

SGML文書は、文書の内部構造の規定である文書型定義“DTD(Document Type Definition)”を持ち、文書データの内部のデータ構造の共通化、標準化を図ることができる(図2参照)。

したがってDocumentBrokerでは、SGML文書を管理し、文書構造を特定した検索機能や部分取り出し機能を実現するために、SGML文書そのものの解析とそれに基づく管理機能を備えている。これらの機能は、基本的に

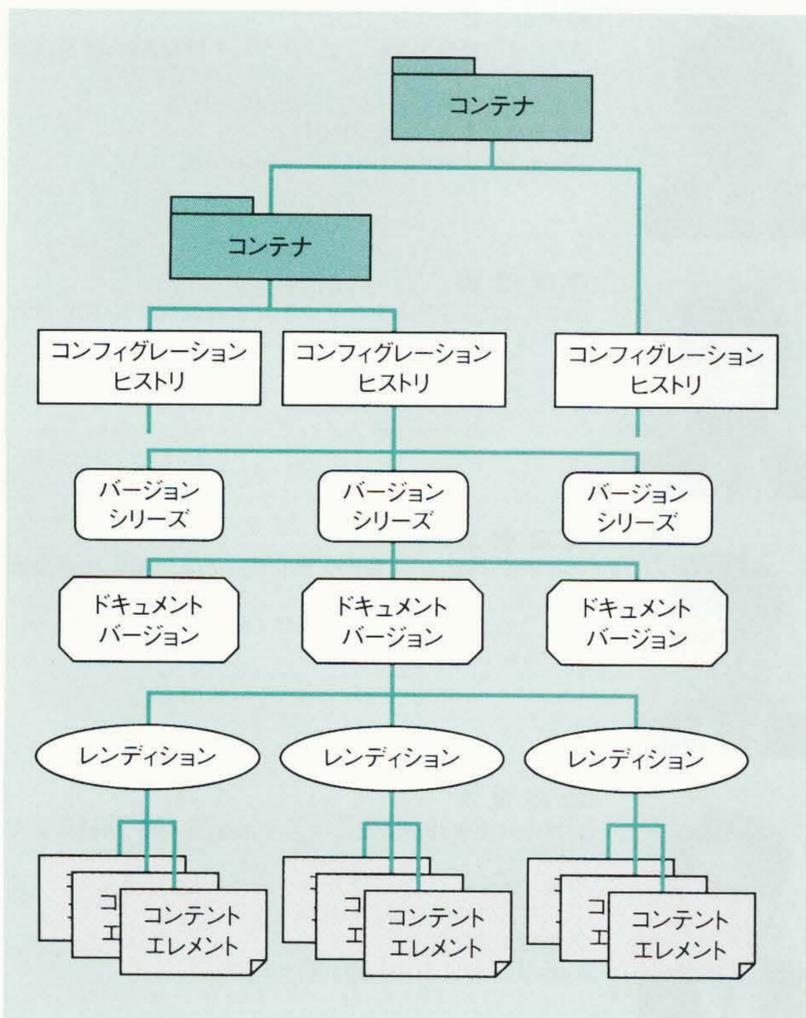


図1 DMA(Document Management Alliance)のオブジェクトモデル

DMAのオブジェクトモデルは、文書を管理する五つの階層から成る。

表1 DMA基本クラスの役割

一つの論理的文書に対して、複数のバージョン系列、複数のコンテンツ形式、および複数の構成要素(ファイル)を管理することができるモデルである。

クラスの名称	役割
コンテナ	文書・コンテナオブジェクトを内部に格納
コンフィグレーションヒストリ	複数のバージョン系列を管理する文書オブジェクト
バージョンシリーズ	一つの系列に属する複数のバージョン系列を管理する文書オブジェクト
ドキュメントバージョン	あるバージョンの文書オブジェクト
レンディション	データの表現形式を管理
コンテンツエレメント	文書オブジェクトの構成要素を管理

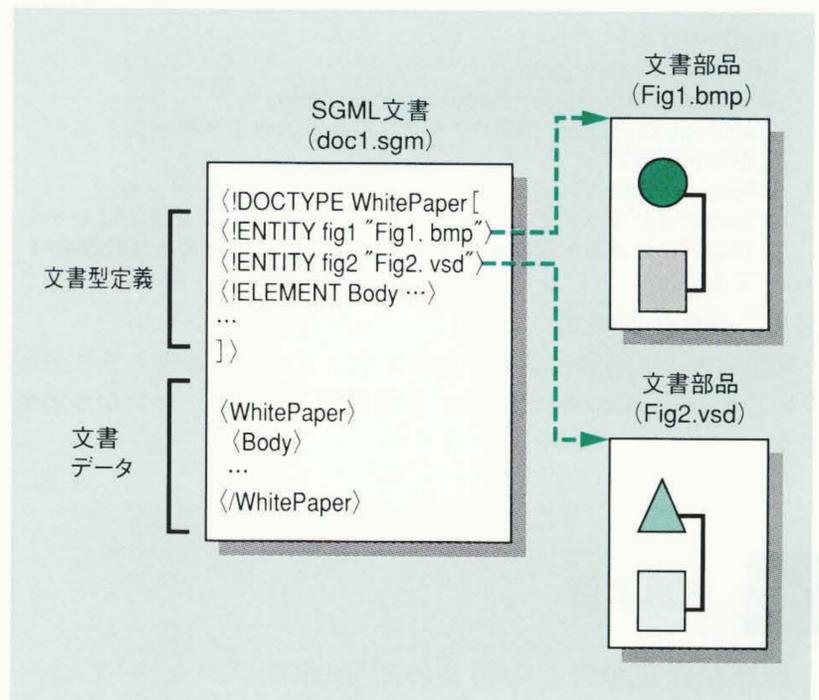


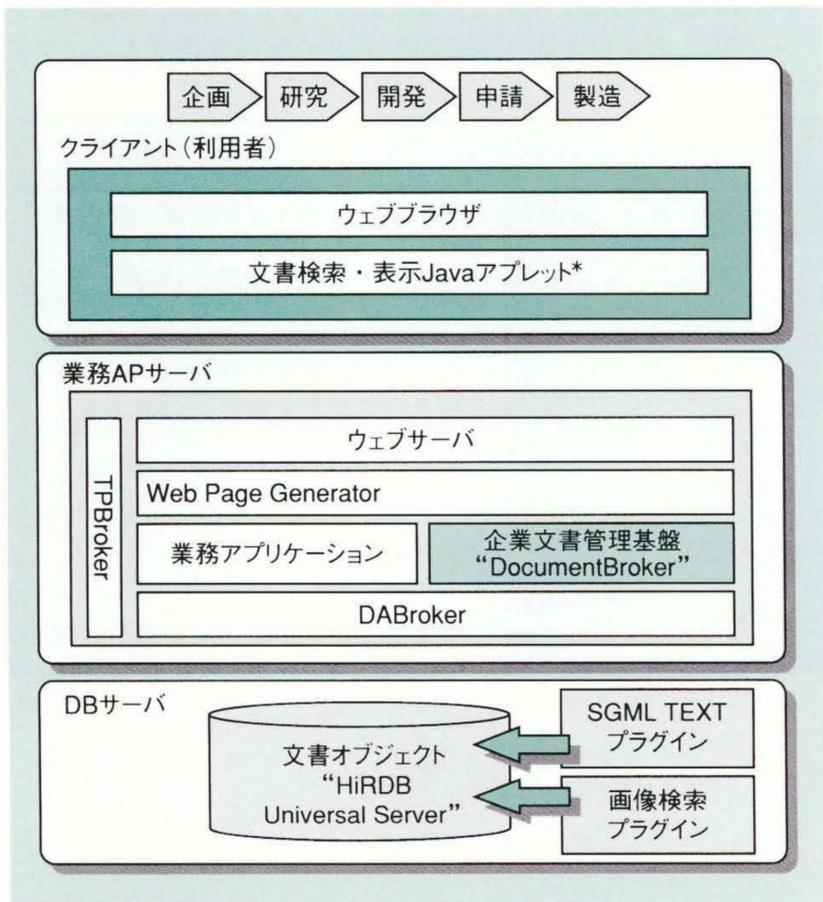
図2 SGML文書のモデル

SGML文書は文書の内部構造を規定する文書型宣言を持ち、文書データの内部のデータ構造の共通化・標準化を図ることができる。

は、ORDBである“HiRDB Universal Server”の「SGML TEXTプラグイン」を利用する形で実現している。

### 3.4 システム構成

DocumentBrokerを使ったウェブ環境でのシステム構成の例を図3に示す。分散オブジェクト下の業務システム構築に際し、文書オブジェクトを管理、操作するためのミドルウェアとして、DocumentBrokerを業務アプリケーションから使用する。



注：略語説明ほか  
 AP (Application Program)  
 TPBroker [ORB (Object Request Broker) 機能]  
 Web Page Generator (DBアクセス結果などのウェブページ整形処理機能)  
 DABroker (分散オブジェクトデータベース アクセス ツール)  
 \*JavaおよびすべてのJava関連の商標およびロゴは、米国およびその他の国における米国Sun Microsystems, Inc.の商標または登録商標である。

図3 システム概念

業務システム構築に際し、文書オブジェクトを管理するミドルウェアとして、“DocumentBroker”を業務アプリケーションから使用する。

4 適用事例

医薬品産業分野での新薬の開発(創薬プロセス)では、企画から始まり、約10年にも及ぶ研究、臨床試験を経て新薬申請に至る。申請に際しては、この間のさまざまなプロセスの成果と膨大な資料を、できるだけ短期間に効率よく申請書類としてまとめあげることが要求される。米国では新薬申請はすでに電子化されており、わが国でも早晚、電子化申請の導入が予想される。

こうした申請業務を支援する業務システムとして、全体としての申請書類を構成する個々の文書の作成や改訂の履歴と、その基になった報告書や実験データなどの関連など、開発過程での業務成果を、長期間にわたって

体系的にかつ着実に蓄積、管理する創薬管理システムの検討が進められている。

5 おわりに

ここでは、企業文書管理基盤“DocumentBroker”について述べた。

文書情報自体の利用拡大に向けた構造化文書、特にXML (Extensible Markup Language)<sup>2)</sup>の標準化の急速な進展は、次世代に向けた文書管理基盤にとって大きな環境の変化といえる。

日立製作所は、こうした変化を注視しながら、DocumentBrokerでの構造化文書対応、全文検索対応の仕様を標準仕様の拡張としてDMAに提案し、標準化に努めるとともに、実システムでの適用事例を基に、各業種対応の共通のノウハウの蓄積と部品群の整備を進めていく考えである。

参考文献・情報

- 1) DMA 1.0 Specification, 1998 (<http://www.aiim.org/dma/dma10/index.htm>)
- 2) 村田：XML入門，日本経済新聞社(1998)

執筆者紹介

**和歌山 哲**  
 1979年日立製作所入社，ソフトウェア開発本部 関西設計部 所属  
 現在，文書管理システムの研究開発に従事  
 情報処理学会会員，IEEE会員  
 E-mail：wakayast@soft.hitachi.co.jp

**伊藤 泰樹**  
 1985年日立製作所入社，ソフトウェア開発本部 関西設計部 所属  
 現在，文書管理システムの開発に従事  
 情報処理学会会員

**福田 俊彦**  
 1989年日立製作所入社，ソフトウェア開発本部 関西設計部 所属  
 現在，文書管理システムの開発に従事  
 情報処理学会会員

**加藤 寛次**  
 1974年日立製作所入社，システム開発本部 開発第1部 所属  
 現在，マルチメディアデータベースおよびドキュメントミドルウェアの研究開発に従事  
 情報処理学会会員  
 E-mail：kanji@isrd.hitachi.co.jp