

ドキュメント管理システムは、従来の専用機からクライアント/サーバ型のシステムに移行している。また、CD-ROM/MO (光磁気ディスク)/DVD に代表される光メディアの進歩により、低価格で大容量のストレージが実現できるようになった。このため、ドキュメントの内容は、従来のテキストやイメージに加え、音声・動画などのマルチメディアデータが含まれるようになりつつある。

当社では、このような新しい潮流を踏まえた、ドキュメント管理技術に関する基本的な要素として、ドキュメントの入力、蓄積、検索、変換、表示、管理などの機能の開発に取り組んでいる。

The market trend of document management systems has moved from conventional proprietary machines to client-server systems. With the great progress made in optical media such as CD-ROM, MO and DVD, mass storage is available at low cost. Moreover, the contents of documents have changed from text to multimedia data including audio/video information.

This paper describes Toshiba's technologies for document management systems.

1 まえがき

わが国のドキュメント管理システムの市場動向は、従来、TOSFile_{TM}に代表される電子ファイリングシステム専用機が中心であった。しかし、最近では、オープン化/ダウンサイジングの流れを受け、汎(はん)用プラットフォームを用いたシステムに移行している。また、ネットワークの普及とそれに伴う情報量の増大により、入力・蓄積・検索・変換・出力といったドキュメント管理を、統合的に行う管理システムの市場は今後ますます拡大していくと考えられる。

一方、ドキュメントの内容も従来の文書・図面に加え、音声・動画などのマルチメディアデータを含むものが増え、これらを効率よく処理する必要がある。

このような、市場動向に対する当社の取組みと、独自に開発した技術を紹介する。

2 ドキュメント管理の現状と将来像

2.1 現状

現在のドキュメント管理システムの対象は主にテキストや図面のイメージである。これらを電子的に保管することで、紙やマイクロフィルムで保管するより、はるかに柔軟な取扱いが可能となる。しかし、このような電子的に保管されているドキュメントは全体の10%に満たない(図1)。

近年、パソコン(PC)の普及に伴い、オフィスにおけるドキュメントもデジタルデータで管理する機会が増加しており、電子化されたドキュメントの比率は高まっているが、

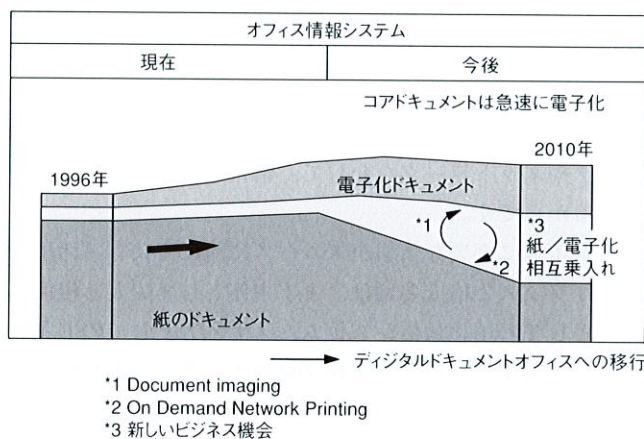


図1. ドキュメント管理システムのビジネスモデル 2000年頃までは紙が増えるが、その後は急速に電子化が進む。

Business model of document management systems

その方向への流れは緩やかである。その理由は、紙の長所を凌駕(が)する電子化ドキュメント媒体が、いまだに存在しないことにもよる。

一方、ドキュメントは従来のテキストから、高度なグラフィックデータやカラー画像などを含む複合ドキュメントに変化しており、これらを効率よく管理することが可能なシステムが求められている。

2.2 将来像

ドキュメント管理で、蓄積・検索のような機能は電子化ドキュメントによるほうが、紙よりもはるかに容易に実現することができる。一方、ドキュメントの表示媒体として

は、取扱いの容易さや表現力など、紙のほうが優位な部分が多い。

このような、紙と電子化ドキュメントのそれぞれの特性を生かした、相互変換機能をもったドキュメント管理システムを、ワークフローやグループウェアに代表されるビジネスプロセスの中に配置すれば、紙という既存メディアの良さと、新しいデジタルメディアの良さを相互に生かした、生産的なオフィス情報システムを構築できると考えられる(図2)。

さらに、効率的で正確な相互変換機能は、オフィスからの保存媒体としての紙を一掃する可能性をもっている。

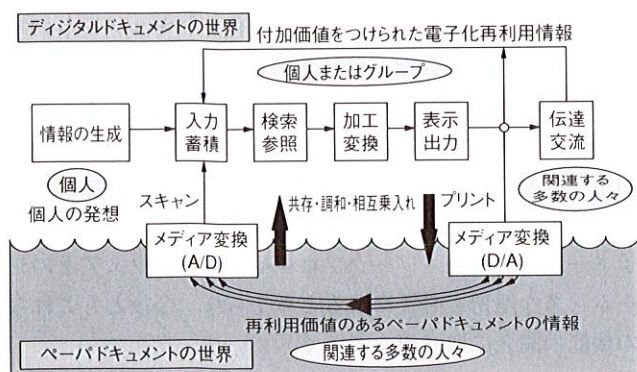


図2. ドキュメント管理システムの将来像 新しいドキュメント管理システムのコンセプト。紙と電子化ドキュメントの相互変換を特長とする。

Concept of document management system in near future

3 ドキュメント管理システムを支える技術

ドキュメント管理の基本的な要素、入力・蓄積・検索・表示のための構成を図3に示す。

3.1 ドキュメントの入力・変換

3.1.1 高速スキャナ ドキュメント管理には、紙ベースのドキュメントを計算機で処理可能な形式に変換するためにスキャナ装置が使用される。PC用のイメージスキャナ装置は、低価格のものが普及しているが、読取り速度は毎分数枚程度で遅い。

一方、業務用として帳票の高速読取りが可能なスキャナの需要は高く、当社では毎分60枚の読取りが可能なOCR(Optical Character Reader)スキャナを商品化している。

3.1.2 ドキュメント変換 CALSにおけるドキュメントの標準としてSGML(Standard Generalized Markup Language)が注目を集めている。SGMLは文書構造に着目したドキュメントフォーマットで、文書の再利用が容易な構造となっている。ワープロやDTP(Desk Top Publishing)がこのSGML対応することで、新規に作成する文書は標準フ

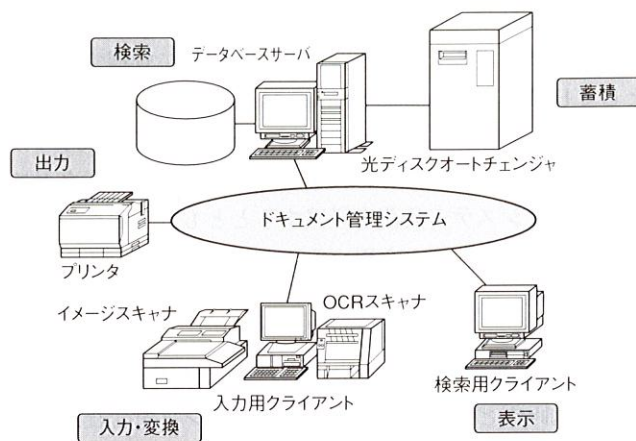


図3. 代表的なドキュメント管理システムの構成例 入力・変換・蓄積・検索・表示・出力機器で構成される。

Typical system configuration

ォーマットに従って生成することが可能になっているが、すでに大量に存在している紙ベースの文書の電子化の手段が必要である。当社では、OCR技術を応用し、スキャナで入力したイメージから、文字認識だけでなく自動的にレイアウト解析を行い文書構造を理解するドキュメントリーダを商品化している。この製品は、文字領域/図や写真などの領域を自動的に切り分ける機能をもっており、従来のテキスト形式のほかに表やイメージを含む複合ドキュメントを、HTML(Hyper Text Markup Language)形式の文書へ変換することが可能である。

3.2 ドキュメントの蓄積

3.2.1 光ディスクの動向 従来、電子ファイリングシステムでは、その容量と信頼性のため、追記型光ディスクがデータの蓄積用の媒体として使用されることが多かった。しかし、書換え可能な光磁気ディスクの大容量化に伴い、追記型光ディスクは限られた分野での利用へと姿を変えている。光磁気ディスクは、標準化により電子ファイリングのデータの蓄積媒体としての普及が進んでいる。

一方、PCの普及に伴い、ドキュメントの配布用の媒体として、CD-ROMが利用されることが多くなっている。

3.2.2 DVDへの進化 1996年にDVD-ROMが製品化されたが、この媒体は、1枚当たりの記憶容量がCD-ROM(650 Mバイト)の約7倍の4.7 Gバイトと大きいため、今後、CD-ROMに変わって配布用の媒体の主力になると考えられる。

一方、書換え可能なDVD-RAMが商品化されると、容量/コストで光磁気ディスクより優位になり、ドキュメントの蓄積媒体の主力になると考えられる。

3.2.3 大容量ストレージ管理 クライアント/サーバ型のドキュメント管理システムで、サーバに格納されるドキュメントの容量は、年々大きくなっている。これに対応

するため、光ディスクオートチェンジャが用いられるが、これを磁気ディスクと組み合わせて効率的に管理する手法として HSM (Hierarchical Storage Management: 階層記憶管理) がある。当社では、従来ソフトウェアにより実現していたこの機能をハードウェアにより実現し、データアクセスの高速化とシステム負荷を軽減するとともに、独自のデータ再配置アルゴリズムを用いることで、より高いスループットを実現している。

3.3 ドキュメントの表示

3.3.1 マルチメディアコンテンツへの対応 最近、音声や動画などのマルチメディアコンテンツを含む複合ドキュメントが作成されることが多くなっている。また、DVD の登場により、従来用いられていた動画圧縮方式の MPEG1 (Moving Picture Experts Group-1) より高品位な MPEG2 が普及し始めており、これらの動画を処理可能な機構が必要になっている。MPEG1 はソフトウェアによるデコードが主流であるが、より高度な処理が必要な MPEG2 は専用ハードウェアを用いる方法が主流で、当社は、PCI (Peripheral Component Interconnect) バスに接続する MPEG2 デコーダを商品化している。

3.4 ドキュメントの検索・管理

3.4.1 電子ファイリングシステム 電子ファイリングシステムは、スキャナから入力されたイメージに、検索のためのタイトル情報を付加して、管理するシステムである。従来の電子ファイリングには、主に専用機が用いられていた。最近の PC とネットワークの普及に伴い、電子ファイリングのシステムは、専用機からクライアント/サーバ型に移行している。当社では、パーソナルから、企業レベルまでの広範囲なレンジをカバーする電子ファイリングシステムとして、FilingStage_{TM}を商品化している。このシステムの特長は、専用機のもっていた機能をクライアント/サーバ型のシステムで実現したことに加え、標準的な SQL (Structured Query Language) による文書データベースのアクセス機能を提供することで、ユーザアプリケーションとファイリングシステムの連携を可能にしたことである。さらに、WWW (World Wide Web) ブラウザからの検索機能や、ワークフローとの連携機能も備えている。

3.4.2 データベース管理 ドキュメントは、データベース技術を用いて管理される。日付・作成者・ドキュメント番号などの値を、データベースのフィールドに割り当て、実際のファイルの格納場所を示すポイントをレコードに格納する方法が採られている。当社では、このしくみを発展させ、SQL によりポイントが示すファイルの内容をアクセス可能にするソフトウェア DBStage_{TM}を商品化している。

3.4.3 全文検索 入力的高速化により、大容量のドク

ュメントデータを容易に作成できるようになった。反面、必要な情報を即座に検索することが困難になりつつあり、高速で柔軟な検索機能が必要になっている。このような状況に対応して、固定的なキーワードの検索を対象としていた従来の検索システムに比べ、すべての言葉をキーワードとして、高速に全文検索することができるソフトウェア FineSearch_{TM}を商品化している。

4 今後の展望

インターネット/イントラネットの普及に伴い、ドキュメント情報をこれらの環境下で扱う機会が増えてくる。また、今後はモバイル環境からのドキュメントアクセスの機会も増大すると思われ、通信の暗号化、ドキュメントの暗号化と鍵(かぎ)の管理、登録されているドキュメントの改ざん防止など、セキュリティの強化が必須(す)である。

オフィス業務の効率化の観点からは、ドキュメントの再利用を容易にするための文書構造化技術や、正確で効率のよいメディアの変換技術の開発が要求される。さらに、電子メールとの連携、ワークフローやグループウェアとのシームレスな連携も必要になると考えられ、今後ともこれらの機能の開発に取り組んでいく。

5 あとがき

コンピュータと通信の発達をもたらす飛躍的な情報量の増大から、ドキュメント管理システムの市場は、急速に拡大すると予想される。また、マルチメディア化するドキュメントにより、コンピュータと映像・音声機器、通信機器が互いに密接に連携したシステム構築の技術も、ますます高度なものが要求される。さらに、環境への配慮からオフィスにおける紙のドキュメントの削減も急務である。今後は、真に使い勝手の良いシステムを旨とし、開発を進めていく。



梶川 茂司 Shigeji Kajikawa

コンピュータ・通信ネットワーク事業統括部 コンピュータ・通信ネットワーク商品企画担当参事。
周辺機器およびファイリングシステムの商品企画に従事。
Computer & Network Div.



山口 晶嗣 Akitsugu Yamaguchi

コンピュータ・通信ネットワーク事業統括部 コンピュータ・通信ネットワーク商品企画担当主任。
周辺機器およびファイリングシステムの商品企画に従事。
Computer & Network Div.